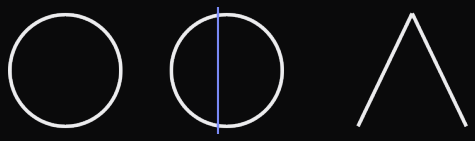

OUTCOME-OWNED
ARCHITECTURE

WHITEPAPER · WORKING DRAFT V1.7

Outcome-Owned Architecture

A new physics of work. Re-deriving the enterprise operating model from the decision up, for the moment the foundation of work itself has changed.

$$\rho = V / (C + R)$$

Jamie Bernard with William Haas Evans

FORESIGHTOPS · FUGUE STRATEGY ADVISORS

Abstract

Enterprise operating models encode assumptions about the physics of work: how fast information moves, what coordination costs, who can execute, and when decisions may be made. Those assumptions, which this paper calls the execution foundation, were fixed at the time each operating model was designed, and most were designed for a foundation that is rapidly becoming unfit for purpose. Capable autonomous agents go beyond automation tooling. They invalidate the assumptions the current model was derived from. The dominant industry response deploys agents into the existing structure of work, which optimizes a shape that is itself the artifact of obsolete constraints, and predictably reproduces the silos, batch boundaries, and coordination overhead of the current organization at higher speed.

This paper articulates Outcome-Owned Architecture (OOA), a sector-agnostic operating model for the enterprise re-derived from first principles when the execution foundation changes, together with its economic engine, the Outcome-Earned Economic Model. OOA is produced by re-deriving the enterprise from the decision up. It takes the recurring decision loop, rather than the process, function, or organizational unit, as its unit of analysis.

It is derived in three movements: an instrumented, empirical trace of a decision's current-state lifecycle; a re-derivation of the same decision under four explicit foundation assumptions, that information is instantaneous, that coordination is largely unimpeded, that execution is autonomous, and that decisions are continuous; and a designed partition of the work into agent-owned judgment and execution loops and human-owned judgment and execution loops, where each agent is bounded by an explicit declaration of what it will not decide. The architecture is completed by a seam register documenting where the foundation has not in fact changed, and by the Outcome-Earned Economic Model, a compute-yield allocation $\rho = V / (C + R)$ that ranks decision loops by value per unit of risk-adjusted cost and derives the workforce partition from where the compute budget line falls.

The paper specifies OOA's constructs, its phased design procedure, governance model, metric system, preconditions, postconditions, and control catalog at a level intended to support conversion into a testable, implementable, and customizable instrument.

Three points of epistemic discipline are stated up front. First, the numerical examples throughout are schema-conformant illustrations, used to demonstrate the constructs and the artifact set. The method's empirical content is carried by the registered hypotheses of Section 11, each currently a prediction. Second, the framework's structural claims are stated as design assumptions the method proceeds under, with the assumptions that carry empirical content restated as falsifiable hypotheses. Third, the method's neutrality with respect to platform selection is treated as an ordering condition the procedure depends on. Section 1.2 states the condition, the mechanism by which it fails, and the controls that detect the failure.

Evidence status and how to read the figures

This is a working draft of a method. Every quantity in the paper should be read as illustrative. This statement licenses the illustrations downstream and will not repeat. The method's empirical content is carried entirely by the hypotheses of Section 11, each stated with a measure and a falsifier, and each currently a registered prediction. The cross-engagement corpus on which those hypotheses will be tested is at the beginning of its accumulation.

A synthetic corpus exists and is described where its results are used. It verifies that the machinery of the economic model behaves correctly under varied inputs. It is not evidence about business models, it is not drawn from engagements, and no figure in it is client-facing. Machinery verification and empirical support are different claims, and only the first is currently in hand.

Where the body of the paper states a regularity, about foundation mismatch, about the persistence of inherited structures, about the partition of human judgment into a small number of classes, it is stating a design assumption the method proceeds under. Assumptions that carry empirical content are flagged at the point of statement and restated as hypotheses (H1 to H6) in Section 11.

A reader evaluating the framework should treat the constructs and the procedure as the object under review, the illustrations as worked examples that demonstrate the procedure's outputs, and the hypotheses as the claims that will eventually be confirmed or refuted by evidence the framework does not yet possess.

Introduction

> *The configuration is not malfunctioning. It is operating as designed.*
>
> OOA V1.7, SECTION 1.1

1.1 Problem statement

An operating model can be read as a set of answers to four questions. How fast does information move? What does coordination cost? Who is permitted to execute? When may decisions be made? Many enterprise cadences are not intrinsic to the underlying work. Each is an accommodation to a foundation in which information arrived in batches, coordination required co-presence or correspondence, execution required human action, and decisions were therefore rationed onto calendars and queues.

When that foundation changes, when state is observable continuously, when machine-to-machine coordination is effectively unimpeded, when software agents execute commitments without human mediation, and when decisions can be evaluated at the cadence of events unbounded from time, the inherited operating model does not simply become less efficient. It becomes unfit for purpose. The structures that compensated for the prior foundation lose the function that justified them while retaining their constituencies, metrics, cadences, and identities. Incremental improvement of the inherited shape optimizes those structures in place. It does not remove them. This is the regularity the method is built around, and it is stated here as a design assumption.

A trace of a single recurring decision through a contemporary enterprise makes the mismatch observable. The working paper illustrates this with one replenishment decision through a consumer-goods manufacturer's regional distribution center: the elapsed time, the count of systems of record, the humans on the critical path, the manual context transfers between systems, the rate of manual override against system recommendations, and a terminal state in which a stakeholder asks "where is my order?" and it is unlikely that a single party in the chain can answer from system state. The figures in that illustration are constructed to the instrumentation specification of Section 4.1. The interpretation the trace is built to surface is the paper's premise in miniature: the configuration is not malfunctioning. It is operating as designed.

The pattern to watch is the response that embeds agents into the inherited flow. That response treats agents as faster humans and places them into a shape derived from human constraints. Conway's Law generalizes the predicted result: systems tend to mirror the communication structures of the organizations that build them, so agents deployed along the existing organization chart tend to reproduce siloed work as siloed automation.

The tooling improves, but the shape persists. The problem this paper addresses is therefore not how to deploy agents. That is a well-documented, evolving, and often-debated set of mechanisms. The problem addressed is this: by what method does an enterprise re-derive the shape of its work when the foundation the current shape assumed has changed, and by what governance, metrics, and controls is the re-derived shape made accountable, reversible, and falsifiable?

1.2 The ordering condition: decision before platform

Platform capability is admissible as a derived output of the method, downstream of the decision inventory. It is inadmissible as an input that shapes which loops are considered or how they are valued.

The method's outputs include conclusions adverse to increased platform consumption: that a given decision loop should not be automated, that a workload does not justify its compute, that the correct design uses fewer licenses or eliminates a loop entirely. The procedure can reach these conclusions only if the decision inventory is established before any platform capability enters the analysis. This is the ordering condition.

The condition generalizes beyond platforms. Any inherited enterprise taxonomy carries the same hazard, because each was drawn under the prior physics of work and each encodes the shape that physics produced. The published business capability model, the process reference model, the application portfolio, and the organization chart all qualify. Admitted upstream of the decision inventory, any of them bounds which loops are considered and the negative conclusions become unreachable by the order in which the work was done. Platform selection is the instance where the incentive to invert the ordering is sharpest, and the capability model is the instance where the inversion is hardest to see, because it arrives with the authority of the enterprise's own architecture function. The rule is the same for all of them, and it is symmetric: admissible as a derived output, inadmissible as a scoping input.

The condition is stated explicitly because it is where the method is most exposed in practice. The ordering can unintentionally invert when the analysis is conducted by a party whose economics depend on the consumption of a particular platform: the platform's capability set bounds the decision inventory, the value attribution tilts toward loops the platform serves well, and the negative conclusions become unreachable. The method does not assume this risk is absent, and

it does not adjudicate whether a given adopter's incentives are aligned. That question is outside its scope. Its responsibility is narrower and dischargeable: to make the ordering condition explicit, to instrument it, and to register its violation as a finding.

The condition is held in place structurally, in three locations. The analysis ordering itself (Section 3, A6), under which platform capability appears downstream of the decision inventory, so the platform answer is derived and stays out of the premise. The admissibility of negative conclusions in the yield allocation and the judgment taxonomy (Sections 4.4, 4.6), exercised at least once per engagement through the mandatory inverse test (Section 10.1), which requires that some loop be carried through the full method with an expected outcome of human ownership or elimination. And the qualification registers, the seam register (4.5) and the value-attribution discipline (7.3), which require explicit documentation of where the design does not work and where value cannot be attributed.

These mechanisms are properties of the procedure and are decoupled from the party running it. An organization whose incentives align with the ordering condition satisfies it with little friction. An organization whose incentives cut against it must rely on the structural controls to hold the ordering in place, and the controls are specified (Section 10) so that a failure to hold it is detectable.

An engagement whose ρ ledger funds nearly every loop it touches, or whose inverse test produces no human-owned or eliminated loop, has not satisfied the condition regardless of intent. The framework's claim is therefore not that any class of practitioner is disqualified from operating-model work. It is that operating-model re-derivation is sound only when the decision inventory precedes platform selection, and that the soundness is checkable from the artifacts.

1.3 Contribution and scope

This paper contributes: (i) a definitional system for the framework's constructs, decision loop, execution foundation, foundation shift, judgment boundary, human-loop taxonomy, seam, work capability, capability anchor, and compute yield; (ii) a basis stated as design assumptions, each carrying empirical content restated as a falsifiable hypothesis in Section 11; (iii) a phased methodology with evidence standards for each phase; (iv) a governance model assigning decision rights, accountability, and change control over the agent layer; (v) a four-family metric system with named anti-metrics and Goodhart safeguards; (vi) explicit preconditions, postconditions, and a three-tier control catalog; and (vii) a validation program structured as testable hypotheses, intended as the bridge from framework to instrument.

The scope is the design of operating models for recurring, instrumentable decision loops within and across enterprise functions. The method is sector-agnostic by construction (Section 3.2) and is illustrated in two domains, supply-chain replenishment and the financial close. Out of scope are

the engineering of individual agents; the selection among specific agent platforms (a derived output under the ordering condition of 1.2); and one-off, non-recurring strategic decisions, which lack the loop structure the method requires.

Intellectual lineage and related work

> *Hierarchies reject a larger proportion of good projects.*
>
> SAH & STIGLITZ, AMERICAN ECONOMIC REVIEW, 1986

OOA is not without ancestry. Locating it serves two purposes: intellectual honesty about what is borrowed, and clarity about what is new.

The method is grounded most deeply in the Theory of Constraints (Goldratt, 1984). From it the method inherits its foundational move: the binding constraint, not the work, is the object of management, and improving anything other than the constraint is local optimization that leaves total throughput unchanged. In a knowledge enterprise the binding constraint is the decision queue. Work waits at the points where a commitment must be made and an authority must act, and the lead time an enterprise actually experiences is dominated by that waiting rather than by the work itself. The seam, in this paper's vocabulary, is where the queue forms: the interface at which a decision waits on a batch boundary, a calendar, or a human's availability while the information required to decide already exists.

Three of the method's load-bearing structures are constraint logic applied to the decision queue. The re-derivation locates the constraint by tracing where the decision waits rather than where the work runs. The insight that human ratification becomes the new bottleneck, once agents take the knowable work, is the five focusing steps generalized from the factory to the seam: identify the constraint, exploit it by sizing review to its real demand, subordinate the agentic flow to it so knowable work paces to the seam rather than the reverse, and elevate it only when it is genuinely the limit.

The compute-yield allocation is throughput accounting on the decision portfolio: loops are funded against a bounded budget in descending yield, the budget line is the funding cutoff, and the shadow price of the best unfunded loop states what one more unit of budget would buy. One caution travels with that reading and is developed in 4.6 and Appendix D.4. Ranking by yield alone does not guarantee that the funded set relieves the binding constraint. Yield measures value per unit of risk-adjusted cost, and a loop can score well while sitting nowhere near the queue that

governs throughput. Under Goldratt that combination buys no throughput at all, which is why the ledger carries a constraint flag beside the rank and hands the conflict to a person rather than resolving it arithmetically.

Business Process Reengineering (Hammer & Champy, 1993) is the closest precedent for the re-derivation move. BPR's injunction, "don't automate, obliterate," anticipated that information technology invalidates rather than accelerates inherited process shapes. OOA departs from BPR in three respects. First, unit of analysis: BPR redesigned processes; OOA re-derives decision loops, a distinction developed in Section 3.2 with material consequences. Second, the human residual: BPR treated humans principally as cost to be engineered out, and its implementation record reflects the organizational resistance that posture provoked; OOA treats human judgment loops as designed first-class outputs, not residue. Third, falsifiability: BPR offered no instrumented baseline and no testable predictions; OOA's trace is a measurement protocol and its collapse claims are stated as bounded, pre-registered predictions.

Lean and value-stream mapping contribute the discipline of walking the actual flow and naming waste. The trace's swivel and stall constructs are recognizable relatives of motion- and waiting-waste. The distinguishing move is again the unit: a value-stream map follows material and information through a process; the OOA trace follows a decision through its lifecycle, trigger, information assembly, commitment, execution, feedback, and instruments the judgment events specifically: overrides, escalations, consensus rituals, and the diffusion of accountability.

Lean asks where the flow waits; OOA additionally asks who decided, on what authority, against what recommendation, and what is generated as a result. The Theory of Constraints sharpens the question to a single point, asking which wait is the binding one, because relieving any other wait moves no throughput. The method takes that sharpening as its starting discipline.

The flow economics underneath the wait claim are drawn from Reinertsen (2009) on queueing and the cost of delay in product-development flow, and from Modig and Åhlström (2012) on the distinction between resource efficiency and flow efficiency. Both supply the same regularity the trace measures directly: in knowledge work the share of lead time spent waiting dwarfs the share spent in value-adding touch. The method treats that regularity as something to measure per loop, not to assume.

Sociotechnical systems theory (Trist and Bamforth, 1951, and the tradition that followed) supplies the second pillar. Its principle is joint optimization: the technical system and the social system must be designed together, and a design that optimizes one against the other tends to fail in practice even when it succeeds on paper. Where the Theory of Constraints governs the flow and the allocation, sociotechnical systems theory governs the human half of the design. Axiom A4, that judgment is designed rather than residual, is the joint-optimization principle stated as a

design rule. The method's most-feared failure mode, judgment laundering, is in these terms a failure of joint optimization: a design that engineered the technical system and left the social system to ossify into rubber-stamping.

Decision-rights economics (Jensen & Meckling on specific versus general knowledge, 1992) supplies the principle that decision authority should be collocated with the knowledge required to exercise it. OOA operationalizes this: agents receive authority over decisions whose knowledge requirements are fully instrumentable; humans retain authority where the requisite knowledge is relational, contextual, or constitutive of the policy itself (Section 4.4).

Naturalistic decision research supplies the positive account of human judgment the method depends on (Klein, 1998, 2008). Field studies of firefighters, intensive-care nurses, and military commanders found that experienced practitioners rarely generate and compare option sets. They match a situation to a repertoire of patterns built through experience, then mentally simulate the first workable option to see whether it plays out. The recognition-primed model is therefore a blend of pattern recognition and deliberate analysis, and the analytical half is what the judgment boundary is built to serve: the agent hands up the relevant quantities with options, and the human owner simulates.

Klein's later work draws the same line from the other side, separating well-ordered decisions whose cue structure is stable and whose feedback is prompt, which reward proceduralization, from ambiguous and time-pressured decisions that defeat rule-based approaches (Klein, 2009). Both schools of decision psychology later specified that boundary jointly (Kahneman & Klein, 2009), and Section 4.4 states where OOA draws it.

Boyd's OODA construct supplies the language of continuous decision loops and the competitive logic of loop tempo. The fourth foundation shift, decisions are continuous, not calendar-bound, generalizes OODA's premise from adversarial maneuver to enterprise operations.

Model-risk management (in the lineage of the Federal Reserve's SR 11-7 guidance, 2011) supplies the governance grammar for systems that decide: validation before deployment, ongoing monitoring, drift detection, documented limitations, and effective challenge. OOA's run-time control catalog (Section 10) adapts this grammar from statistical models to autonomous agents.

THE NOVELTY CLAIM, BOUNDED

What OOA combines, and what no single constituent tradition contains, is: the decision loop as unit of analysis; the re-derivation as a formal design operation; the judgment boundary as a mandatory, falsifiable artifact; a taxonomy of human loops claimed exhaustive and testable as such; the seam register as institutionalized accounting for partial foundations; and compute yield as the allocation mechanism from which workforce definition is derived. The novelty claim is itself bounded: it asserts a novel combination and is exposed to test by H1.

Theoretical basis

> *Organizations tend to ship their constraints.*
>
> 00A V1.7, SECTION 3.1, A2

3.1 Design assumptions

The framework proceeds from six assumptions. Four of the six (A1, A3, A4, A5) carry empirical content and are restated as testable hypotheses in Section 11. The method is constructed so that if an assumption fails for a given enterprise, the failure surfaces as a finding. Each assumption below is paired with the posture the method takes toward it and, where applicable, the hypothesis that exposes it to refutation.

A1 · WORK DECOMPOSES INTO DECISIONS AND EXECUTION

Enterprise work decomposes, to a first approximation, into recurring decision loops, commitments of resources under uncertainty with a trigger, an information set, a commitment act, and a feedback path, and execution in service of commitments already made. Activities that appear to be neither, meetings, reconciliations, status reports, swivel-chair data transfer, are on inspection most often coordination overhead incurred to assemble the information set or to diffuse accountability for a commitment, and are treated as properties of decision loops.

> **Posture.** The decomposition is an idealization, and the method does not claim it is exhaustive without remainder. Some of what coordination rituals accomplish is trust maintenance and shared sense-making, which the decision lens underweights. The Meaning Loop (4.4) absorbs part of this, and Section 12.2 records the residual honestly. A1 is used as a working decomposition and can be overridden when minimally biased human judgment deems it necessary.

A2 · OPERATING MODELS ARE FOUNDATION-SHAPED

The structure of an operating model, its cadences, hierarchies, handoffs, and rituals, is treated as a rational derivation from the execution foundation assumed at design time. This is Conway's Law generalized from software architecture to the architecture of work: organizations tend to ship their constraints.

> **Posture.** Stated as the framing assumption that motivates re-derivation. It is not independently tested as a hypothesis. Its consequence, that retrofit reproduces inherited shape, is the testable claim, carried by H2.

A3 · QUALITATIVE FOUNDATION CHANGE FAVORS RE-DERIVATION OVER RETROFIT

When foundation assumptions change qualitatively, batch to instantaneous, mediated to unimpeded, human-required to autonomous, calendar-bound to continuous, the design assumption is that the optimal shape of work changes discontinuously, such that incremental improvement of the inherited shape converges on a local optimum of an obsolete landscape. The mechanism is stated in 3.4.

> **Posture.** This is the framework's most consequential empirical assumption and is exposed to test by H2. If retrofit reliably reaches outcomes indistinguishable from re-derivation, A3 is weakened.

A4 · JUDGMENT IS DESIGNED INTENTIONALLY

The human loops in a re-derived model are positively specified and are not expected to be in a liminal state of whatever the agents cannot yet do. They are loops whose decision variables are non-instrumentable, non-optimizable, or constitutive of the policy that governs the agents themselves. The assumption is that a design which defines human work as the automation residual will mis-assign human judgment, produce accountability vacuums, and tend to decay into rubber-stamping (the judgment-laundering failure mode, Section 12).

> **Posture.** A4 governs the method's human-side discipline and is exposed to test by H3 and by the laundering health metrics (7.4). A recurring class of human judgment that resists positive specification would qualify A4 and is treated as the most valuable negative result the program can produce.

A5 · WORKFORCE DESIGN IS A CONSTRAINED ALLOCATION PROBLEM

Agentic capacity is bounded by compute, which carries real cost. Each candidate agentic loop *i* has a compute yield $\rho = V / (C + R)$, value delivered per unit of risk-adjusted cost, where the denominator carries both what the loop costs to run and what it costs to be wrong. The assumption is that rational design ranks loops by ρ, funds them in descending order until the compute budget line, and derives the workforce from where that line lands, subject to the judgment constraints of A4 (which dominate ρ: a loop assigned to humans by the taxonomy is not a funding candidate regardless of its yield).

> **Posture.** A5 presumes the enterprise can bound and price both cost and risk. Where it cannot price compute, the allocation degenerates (Section 8.4), and ρ flatters every loop whose errors are unpriced. The hard parts empirically are the two attributions, V and R, whose hazards are stated in 11.2. The framework's discipline is symmetric: an unattributable value is recorded as unattributable rather than allocated by narrative, and an unpriced risk is surfaced rather than assumed to be zero.

A6 · NEUTRALITY IS A PROPERTY OF ANALYSIS ORDERING

Any method in which platform capability, or any other inherited enterprise taxonomy, appears upstream of the decision inventory inherits that artifact's incentive geometry and forfeits the admissibility of negative conclusions. Neutrality, where the method achieves it, is therefore a property of the order in which analysis is performed (1.2).

> **Posture.** The method claims the condition's satisfaction is checkable from the artifacts, and registers violation as a finding. Stated as a structural condition the method depends on, with the failure mode and detection specified in 1.2 and the controls in 10.1 and 4.6. H4 is its empirical form: a method whose output distribution approaches "agentic everywhere" fails discriminant validity.

3.2 The decision loop as unit of analysis

The choice of unit is the framework's keystone and the source of its sector-agnosticism. Processes, functions, and organizational units are foundation-contaminated objects: their boundaries were drawn by the prior physics of work. "Demand planning" exists as a function in part because forecasting was batch. "The close" exists as a process because the ledger was periodic. Taking these as units of analysis risks smuggling the obsolete foundation into the redesign as a premise.

The enterprise's own published business capability model belongs in that set, and it is the member most likely to be waved through, because it arrives carrying the authority of the architecture function. Its boundaries were drawn under the prior physics like every other inherited taxonomy, and its identifiers commonly encode the organization chart directly. Its licensed use is the capability anchor (4.7), a reference from a loop out to a leaf node of that model, carried so enterprise architecture can place a re-derived design in its portfolio. The anchor reads the model as a destination. It never reads it as a scope. Contamination is a reason to control the direction of the reference, and it is not a reason to refuse the reference.

The decision loop, by contrast, is treated as foundation-invariant. Some quantity of inventory will be committed to some location; some carrier will be tendered some load; some accrual will be booked at some value; some invoice exception will be resolved some way. These commitments exist under any foundation. Only their shape, cadence, information assembly, authority, and

execution path are foundation-dependent. The decision loop is therefore the stable object the method holds fixed while everything around it is allowed to change. The cognitive evidence points at the same grain: expertise fractionates by task rather than by person (Kahneman & Klein, 2009), so a disposition assigned to a role is assigned to a bundle, and the loop is the level at which skill and authority can actually be matched.

> Formally, a decision loop is a five-tuple ⟨τ, I, π, χ, φ⟩: a trigger class τ (the event family that opens the loop), an information set I (the state required to decide), a policy π (the mapping from information to commitment, whether encoded or tacit), a commitment act χ (the resource allocation that closes the loop instance), and a feedback path φ (how outcomes re-enter I and revise π).

THE FIVE-TUPLE

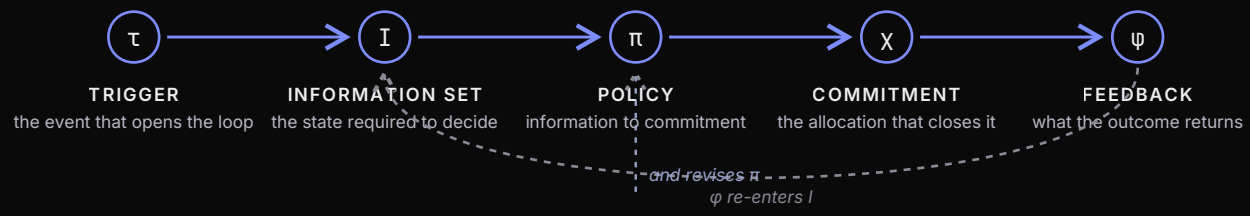

FIG. 3.2 · THE DECISION LOOP AS A FIVE-TUPLE ⟨τ, I, π, χ, φ⟩

Every construct in the framework attaches to this tuple: the trace measures how τ-to-χ traverses systems and humans today; the foundation shifts alter the physics of I, χ, and the loop's cadence; the judgment boundary partitions π; the seam register marks where χ crosses into unshifted territory; and ρ prices the loop.

The sector-agnosticism claim now has precise, falsifiable content: none of the framework's primitives quantifies over any domain-specific type. Triggers, information sets, policies, commitments, feedback, swivels, stalls, overrides, judgment boundaries, seams, and yields are typeless with respect to industry. The supply-chain material in the working papers is payload; the method is carrier. The claim fails if applying the method in a new domain forces the introduction of a new primitive, and is registered as hypothesis H1.

3.3 The execution foundation and the four shifts

The execution foundation is the vector of physical-economic assumptions an operating model makes about work. The framework factors it into four dimensions, each with a legacy value and a shifted value. The shifts are stated as design assumptions adopted for the re-derivation.

TABLE 3.3 · THE FOUR FOUNDATION SHIFTS

DIMENSION	LEGACY VALUE	SHIFTED VALUE (DESIGN ASSUMPTION)
Information latency	State is known in batches: overnight runs, weekly forecasts, bi-weekly reports, monthly actuals.	State is observable continuously. The forecast is the live state of the demand graph. The ledger is a stream rather than a snapshot.
Coordination cost	Alignment requires meetings, emails, escalation chains, consensus rituals. Coordination is expensive and therefore rationed.	Coordination between deciding entities is machine-mediated and effectively unimpeded. Consensus is the ongoing committed state of a shared graph.
Execution agency	Acting on a decision requires human hands: keying, calling, walking, signing.	Software agents execute commitments directly against systems of record. The hand-off is direct and agent-owned, with no human-in-the-loop smokescreen.
Decision cadence	Decisions are calendar-bound and rationed onto the cycle, because the other three dimensions made each decision expensive.	Decisions are continuous, re-evaluated at every material state change, owned by an agent, a human, or both, with policy and flow governing.

The two-step that follows is deliberate. The re-derivation (Section 4.2) assumes all four shifts and derives the ideal shape of work. The seam register (Section 4.5) then reconciles that ideal with the empirical reality that the shifts are partial and uneven. Some may never occur because the value in the shift costs too much, in currency or opportunity. Deriving from the impure foundation contaminates the design with accommodations that harden. Deriving pure and then engineering the seams keeps every accommodation named, owned, and removable when or if the seam closes.

3.4 Discontinuity mechanism

A3 assumes retrofit does not reach the re-derived shape. The mechanism is the framework's answer to "why not just add agents to what we have?" The inherited shape contains structures whose function is to compensate for legacy foundation values and constraints. Under the shifted foundation these structures have no compensating function, but they retain constituencies, metrics, and identities. Local optimization tends to preserve them because each is locally rational.

Only a global re-derivation, which never instantiates these structures in the first place, reliably escapes them. The prediction this yields is concrete and testable: agents bolted onto the inherited flow tend to produce agents that draft the email for the consensus meeting rather than a flow with no meeting, because the retrofit inherited the meeting as a requirement. Whether this prediction stands across deployments is the substance of H2.

Constructs

> *Humans, in the re-derived model, are not in the loops; they own loops.*
>
> OOA V1.7, SECTION 4.4

The framework comprises six constructs and one output artifact set. Together they constitute the operating-model definition: the answer to who decides what, at what cadence, under what policy, with what accountability.

4.1 The empirical trace and its instrumentation ontology

The trace is a measured reconstruction of a decision loop's current-state lifecycle, from trigger to commitment to fulfillment, across every system touched and every human in the chain. It is intended as an empirical protocol whose counts are to be observable, reproducible, and triangulated (methodological standards in Section 5, Phase 1; reliability exposed to test by H5). The instrumentation ontology, the framework's measurement vocabulary, consists of eight constructs, each with a counting rule.

TABLE 4.1 · THE INSTRUMENTATION ONTOLOGY

CONSTRUCT	COUNTING RULE, AND WHAT IT MEASURES
Human-in-chain	Any person whose action or approval is on the critical path from trigger to commitment for the median loop instance. Consulted-but-not-blocking parties are recorded but not counted, so the value is not falsely inflated.
System-of-record touch	Any distinct system in which loop state is created, transformed, or persisted, including shadow systems. Spreadsheets and shared drives are systems of record if commitments depend on them.
Swivel	A manual transfer of loop state between systems: re-keying, copy-paste, export-import, reading one screen while typing into another. Swivels measure the coordination cost humans pay on the foundation's behalf and tend to surface capital waste as a byproduct.
Stall	An interval in which the loop waits on a batch boundary, a calendar, a standing cycle, or a human's availability while the information required to proceed already exists. Stalls measure foundation latency, distinct from genuine information absence (see 11.2).
Override	A human reversal or material modification of a system recommendation. The override rate is the single most diagnostic trace metric: a high rate indicates that encoded policy π and operative policy have diverged, and the real policy lives in tribal knowledge.
Batch boundary	Any point where the loop must wait for a scheduled aggregation.
Consensus ritual	A synchronous gathering whose output is agreement on a number or a state, a structure predicted by legacy coordination cost and whose elimination is predicted by shifted coordination cost.
Accountability diffusion	The terminal trace question: when a stakeholder asks the loop's defining question, can any single human in the chain answer from system state? A "no" is the qualitative signature of a foundation-mismatched model.

INTERPRETIVE RULE

Trace metrics are measurements of the foundation mismatch. A high override rate is the measured distance between encoded and operative policy. The trace's stance, that the configuration is working as designed for a foundation that no longer applies, is a methodological commitment, because traces conducted as performance audits tend to produce defensive data (7.1).

4.2 Re-derivation

Re-derivation is the formal design operation: hold the decision loop's tuple ⟨τ, I, π, χ, φ⟩ fixed, assume the four shifts have occurred, and derive the minimal flow that produces the commitment. The discipline is subtractive: every structure in the derived flow must be forced either by the loop itself or by a declared seam. Nothing is carried over simply because it exists today. The standard of success is that the derived flow could be presented to someone who had never seen the current state and read as obvious. Risk in the re-derivation is inherent because it requires explicit detachment from current sequence constraints.

4.3 The judgment boundary

The judgment boundary is a mandatory, per-agent declaration of three things: what the agent will not decide; the hand-up conditions under which it pages a named human (role or job); and the form of the hand-up, the relevant quantities presented with options.

The boundary is the framework's central safety and accountability artifact, and it is deliberately falsifiable: a design in which any agent's boundary cannot be stated is, by definition, incomplete. The boundary also has a contractual character, as hand-up conditions are defined control events.

Two corollaries govern the boundary's integrity. First, agents do not set their own thresholds: every parameter of every boundary is owned by a human in the Policy Loop (4.4), under change control (Section 6). Second, the boundary is asymmetric by design: agents may always hand up; they may never hand down. No agent may delegate to a human a decision inside its declared scope as a means of diffusing accountability for it.

Hand-ups are typed, and the three types are declared separately because they are governed differently. An **exception hand-up** fires when a declared threshold or condition is met. A **trade-off hand-up** carries a commitment whose decision variable the policy cannot settle. A **signal hand-up** carries something the agent noticed that no declared condition covers: a weak signal, a dissenting reading of the evidence, or the conditional implication of events that are individually unremarkable. Exception and trade-off hand-ups firing outside their declared conditions are control events (10.2).

The signal channel is exempt from that control by construction, because a channel that penalizes unanticipated surfacing is what Klein and Snowden call "organizational policies that filter weak signals", and such policies are the documented mechanism by which groups dismiss what individuals in them have already noticed (Klein, Snowden, & Chew, 2007). The signal channel carries no threshold and confers no authority. It obliges the receiving human loop to record the signal and its disposition.

The decision loop, partitioned by the judgment boundary

A cybernetic control system. Outcomes revise policy; ownership transfers only at the seam.

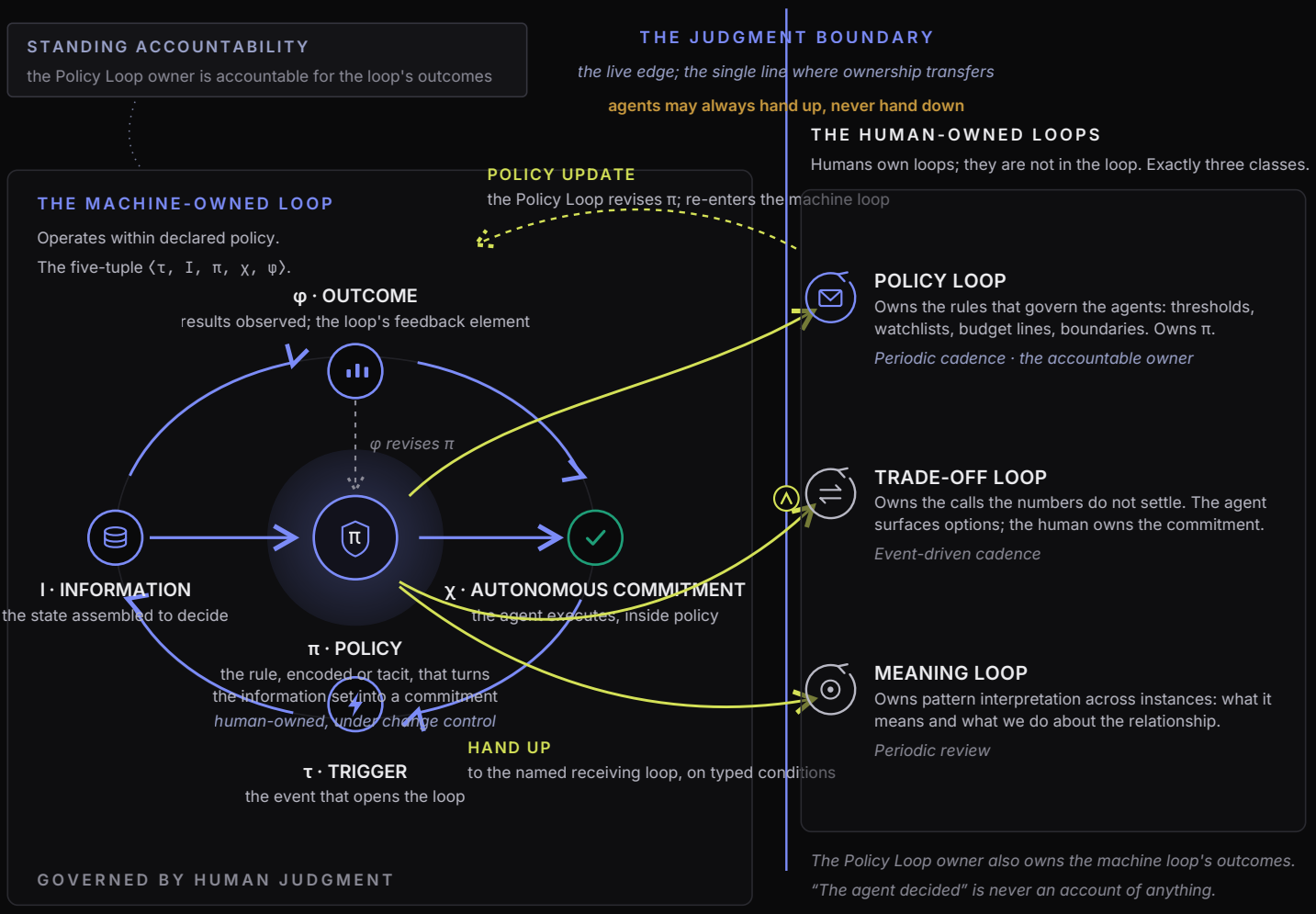

FIG. 4.3 · THE DECISION LOOP, PARTITIONED BY THE JUDGMENT BOUNDARY. A CYBERNETIC CONTROL SYSTEM. THE MACHINE-OWNED LOOP OPERATES WITHIN DECLARED POLICY AS A FIVE-TUPLE. OUTCOMES REVISE POLICY; OWNERSHIP TRANSFERS ONLY AT THE SEAM. AGENTS MAY ALWAYS HAND UP, NEVER HAND DOWN. HUMANS OWN LOOPS; THEY ARE NOT IN THE LOOP.

4.4 The human-loop taxonomy

The framework claims that all designed human judgment in a re-derived model falls into exactly three loop classes. The claim of exhaustiveness is deliberate and testable (H3). It is held as a strong working hypothesis, and the method treats any unclassifiable human judgment as evidence against it. The three classes rest on a positive account of human judgment rather than on a list of what agents cannot yet do.

The Trade-off and Meaning loops are where an experienced person reads a situation as a pattern and commits, then simulates the option before committing. There is a science of the human judgment the method designs for, and these loops are where it lives. That is the recognition-primed judgment documented in naturalistic decision making, summarised as a blend of intuition and analysis (Klein, 2008), in which the simulation carries the analytical weight. That is what A4 means by designed rather than residual.

TABLE 4.4 · THE THREE HUMAN-LOOP CLASSES

LOOP	OWNS	DEFINING QUESTION AND CADENCE
Policy Loop	The rules that govern the agents: thresholds, watchlists, risk postures, budget lines, boundary parameters. Authority over π itself, which is policy-constitutive knowledge.	Which decisions belong to machines and which to humans in a defined governance period? Cadence: periodic. Acts as a standing decision-rights body.
Trade-off Loop	Commitments whose decision variable is non-optimizable. Authority where I is relational, which is Jensen and Meckling specific knowledge.	Which way do we go when the numbers do not settle it? The agent forces the trade-off to the surface with complete information and options. The human owns the call. Cadence: event-driven, defined in policy with a named owner.
Meaning Loop	The interpretation of patterns across loop instances: what the data means for strategy, relationships, and the model itself. Authority over ψ at the population level, which is interpretive knowledge.	What does this pattern mean, and what do we do about the relationship: renegotiate, diversify, exit? Cadence: periodic review.

Humans, in the re-derived model, are not in the loops. They own loops. The taxonomy's theoretical alignment is clean: the Policy Loop holds authority over the policy, the Trade-off Loop where the information is relational and non-instrumentable, and the Meaning Loop over interpretation at the population level. The inversion is the heart of the workforce design.

The partition of work into agent-ownable and human-owned rests on a knowability distinction that Cynefin states cleanly. Where cause and effect are knowable, the clear and complicated domains, a rule can run and an agent can own the loop. Where they are not, the complex and chaotic domains, the response is to sense and probe rather than apply a rule, and the loop stays human. Whether the three classes are in fact exhaustive across real-world deployments is H3, and a recurring coherent fourth class would force a principled extension.

The partition agrees with the conditions for trustworthy intuitive expertise stated jointly by the two rival schools of decision psychology (Kahneman & Klein, 2009): skilled human intuition and executable policy alike require environments of high validity with learnable regularities and prompt feedback, which is where an agent can own the loop. Where validity is low, neither a rule nor an expert's confidence holds predictive authority, and human ownership in that territory rests on accountability, relational knowledge, and non-optimizable decision variables rather than on claimed foresight.

4.5 Seam register

The seam register is the framework's institutionalized accounting for partial foundations. A seam is any interface at which the re-derived flow meets a foundation that has not shifted to a re-derived state. The seam is also where the queue forms in the re-derived flow. It is the constraint the method must then manage, in the constraint-theoretic sense: the point whose throughput governs the throughput of the whole, and the only point where adding capacity changes the system's output.

Each seam entry records: the unshifted dimension; the friction it imposes on the derived flow; the seam pattern applied, encapsulate behind an interface so the integration is achievable when the counterparty shifts, instrument the friction so its cost is continuously measured, or, for regulated seams, switch design patterns entirely (human attestation at the decision point, immutable ledger for the trail, separation of duties for the controls); the named owner; and the closure condition, what change in the world retires the seam.

Two principles govern seams. Seam denial, pretending the counterparty has shifted, is the canonical failure mode of integration programs and is where last-mile integration risks failure. Seams are not defects of the design: a register that says "the regulated fraction follows a different pattern" is evidence of method integrity.

4.6 Compute yield, risk, and the budget line

For each candidate agentic loop *i*, define compute yield $\rho = V / (C + R)$, where V is the value attributable to agentic operation of the loop (Section 7.3 specifies the attribution discipline), C is the fully loaded compute and operations cost of running the loop agentically, and R is the loop's risk cost: the expected cost of the loop's autonomous commitments being wrong, priced against the same trace baseline as V. In short: C is what the loop costs to run; R is what it costs to be wrong, and a loop's yield is charged for both.

R is built from the loop's own decision structure, along two axes the design keeps separate. **Control**: a rule-bounded commitment that errs inside its policy band carries only the in-band cost of that error, net of what is recovered at the next true-up, whereas a genuine unsupervised judgment that errs carries the full cost of the misjudged commitment. **Nature**: a transactional commitment, where the agent acts directly, realizes its error directly, whereas an informational commitment, where the agent informs a human who then acts, realizes its error only through the action-and-recovery chain that follows.

Commitments under human ratification, the seam pattern (4.5) and the regulated edge (6.3), contribute only their residual escape: the wrong calls that survive review. R is therefore the economic expression of the safety architecture the rest of the framework specifies qualitatively. The judgment boundary (4.3), the human-loop partition (4.4), and the regulated-edge pattern (6.3) are the structures that hold R down, and a loop whose boundary is wide, whose human judgment is unsupervised, and whose commitments are irreversible prices that exposure into its yield. The full computation is Appendix D.

The allocation procedure ranks loops by ρ descending, so that a loop is funded on the value it produces per unit of risk-adjusted cost, and a high-value loop whose risk cost is large ranks below a leaner loop it would have outranked on value alone. It applies the judgment constraints (loops assigned to humans by the taxonomy are not candidates regardless of ρ; A4 dominates A5). And it funds down the ranking until the compute budget is exhausted. The budget line derives the workforce: loops above the line are agentic; loops below it remain human or are eliminated where the trace shows they were pure foundation compensation. The shadow price of the best unfunded loop is the marginal value of relaxing the constraint by one unit of budget.

Ranking by yield does not by itself locate the constraint, and the two questions must be kept apart. A loop can carry high value per unit of risk-adjusted cost while sitting nowhere near the queue that governs the enterprise's throughput, and funding it in that position buys speed on a non-constraint, which under the Theory of Constraints buys nothing. The ledger therefore carries a constraint flag beside the rank, recording whether each loop sits at the binding constraint. Where the top-ranked loop relieves no constraint and the loop that does falls below the budget line, the instrument surfaces the conflict and hands it to a named human owner rather than resolving it arithmetically. That is the honest treatment, and Appendix D.4 works the case through.

Workforce definition is thus an output of yield allocation under judgment constraints, not an input from the org chart. This is also the second structural locus of the ordering condition (1.2): ρ-ranking can, and where the evidence warrants will, conclude that some loops do not justify their compute, a conclusion the method must be able to produce for the neutrality property to hold, and whose presence or absence is checkable from the ρ ledger.

4.7 The output artifact set

A completed OOA engagement produces, as its operating-model definition, exactly six artifacts. Postconditions (Section 9) are stated against this set.

1. **Decision inventory**, the loops, as five-tuples, with trace baselines.
2. **Judgment boundary catalog**, every agent's line, hand-up conditions, and owners.
3. **Human-loop registry**, every Policy, Trade-off, and Meaning loop, with named owner, cadence, decision rights, and information rights.
4. **Seam register** (4.5).
5. **The ρ ledger**, the ranking, the budget line, the funding decisions, and the eliminations.
6. **Control catalog** (Section 10), instantiated against the specific design.

Three of the six carry a **capability anchor** as a required attribute: the decision inventory, the human-loop registry, and the control catalog. The anchor is a reference from a loop, or from a work capability beneath it, to a leaf node in the enterprise's own published business capability model, carried together with the owning architecture authority and the model version. It is an attribute rather than a seventh artifact, which preserves both the exactly-six claim and the P-series postconditions stated against it.

The anchor earns its place rather than being tolerated. An enterprise large enough to need this work already has an architecture function and a published capability model, and a re-derived design that cannot be placed in that portfolio is refused on governance grounds before anyone evaluates its merit. The security function needs it for the same structural reason: enterprise control frameworks scope controls to capabilities, so without the anchor every control in Section 10 arrives as net-new to be argued for rather than as an instance of one the enterprise already owns. The direction of the reference is what the method controls, and the anchor direction test (10.1) is where that control lives.

Methodology

> *A phase is complete when its exit criterion is met, not when a calendar interval elapses.*
>
> OOA V1.7, SECTION 5

Seven phases, 0 through 6. Each is specified by purpose, activities, evidence standard, outputs, and exit criterion. A phase is complete when its exit criterion is met, not when a calendar interval elapses. The phases are sequential in logic, but Phases 1 and 2 may iterate per decision loop. The procedure is described in terms of entry and exit conditions rather than durations.

PHASE 0 · SCOPING AND THE DECISION INVENTORY

Purpose. Select the unit or units of analysis and establish the engagement's epistemic ground rules.

Activities. Identify candidate decision loops by their commitments (what resources get allocated, recurringly?). Formalize each candidate as a five-tuple. Select tracing candidates by materiality (commitment value × frequency) and by suspected foundation mismatch (visible batch boundaries, known override folklore, chronic 'where is it?' questions). Establish the front-of-work gate: confirm role boundaries, work-breakdown ownership, and IP-provenance conventions before any artifact is produced, with version history and repository tagging. Record the capability anchor for each loop entered into the inventory, downstream of the commitment that put it there and never as the reason it was considered.

Evidence standard. Every loop in the inventory must be stated as a commitment such that someone outside the function recognizes it as a decision. Loops that can only be stated as process names fail intake.

Exit criterion. A ranked inventory with two to five loops selected for tracing and named data access secured for Phase 1.

PHASE 1 · THE CURRENT-STATE TRACE

Purpose. Produce the empirical baseline.

Activities. Reconstruct the median loop instance and at least one high-percentile instance end to end. Triangulate three evidence classes, system logs and timestamps, direct observation or screen-share walkthroughs, and practitioner interviews, under the rule that counts come from logs and observation while interviews supply only interpretation. Apply the instrumentation ontology (4.1) with its counting rules. Where a counting rule is ambiguous in context, resolve it in writing before counting and apply the resolution uniformly. Trace the exception path as well as the happy path: the override rate in particular cannot be measured from the happy path.

Evidence standard. Every count independently reproducible by a second analyst from the same evidence. Every stall and swivel attributed to a foundation dimension.

Outputs. The trace document in the working-paper grammar: acts, timestamps, systems touched, artifacts, swivels, stalls, overrides, and the cumulative-weight metrics.

Exit criterion. The trace metrics are signed by the practitioners who live the flow ('yes, that is our Tuesday'), practitioner recognition being the trace's validity check, and the loop's defining accountability question has been asked and the answer recorded.

PHASE 2 · RE-DERIVATION

Purpose. Derive the shifted-foundation shape.

Activities. Hold the five-tuple fixed. Assume the four shifts have or will occur. Derive the minimal flow that is unbounded from its current state. For every element of the derived flow, record its forcing reason (forced by the loop, or forced by a declared seam, no third category). Specify each agent by what it perceives, what it commits, which systems it touches, what artifacts it produces, and, mandatorily, its judgment boundary. Apply the subtractive discipline: any structure surviving from the current state must re-justify its existence from the shifted foundation, and consensus rituals, override passes, and batch boundaries are presumptively eliminated.

Evidence standard. No agent can exist without a complete boundary declaration, and the forcing-reason audit admits no derived structure without a recorded force.

Outputs. The re-derived flow, the judgment boundary catalog (draft), and the predicted collapse metrics.

Exit criterion. The collapse metrics are stated as predictions with bounds before any deployment, so that they remain falsifiable (7.2, H2).

PHASE 3 · JUDGMENT PARTITION

Purpose. Design the human side as a first-class output.

Activities. Enumerate every hand-up condition in the boundary catalog and classify each into the three-loop taxonomy. Any hand-up that resists classification is escalated as a potential taxonomy counterexample (recorded for H3) and resolved by either reclassification or explicit extension under change control. Define each human loop as a role specification: owner (named, singular), cadence, decision rights, information rights, and the loop's defining question. Apply the anti-laundering test: for each loop, specify what engaged exercise of the human judgment looks like versus rubber-stamping, and what metric would detect the difference (7.4).

Outputs. The human-loop registry.

Exit criterion. Every hand-up condition in the catalog maps to exactly one registered human loop with a named owner.

PHASE 4 · SEAM ENGINEERING

Purpose. Reconcile the derived ideal with the partial foundation.

Activities. Inventory every interface where the derived flow meets an unshifted counterparty or constraint. For each, select the seam pattern (encapsulate, instrument, or regulated-edge), assign the owner, and state the closure condition. For regulated seams, instantiate the regulated-edge pattern in full: human attestation at the decision point, immutable decision ledger, separation of duties between the agent that proposes and the system that commits.

Evidence standard. No seam may be closed by assumption. A counterparty's foundation status is established by evidence.

Outputs. The seam register.

Exit criterion. The derived flow plus the seam register together account for one hundred percent of the loop's path.

PHASE 5 · YIELD ALLOCATION

Purpose. Derive the workforce.

Activities. Attribute V for each candidate loop against the trace baseline using the value-attribution discipline of Section 7.3. Estimate C fully loaded (compute, integration amortization, run operations, the human loops' time). Build R (expected annual cost of the loop's autonomous commitments being wrong) per loop from its own error structure. Apply the judgment constraints (loops the taxonomy assigns to humans are not funding candidates regardless of ρ, A4 dominates A5). Rank ρ descending, set the budget line, classify every loop as agentic, human, or eliminated, and record every classification's reason. Record the constraint flag beside each rank, and where the funded set relieves no binding constraint, surface that as a finding for the Policy Loop rather than resolving it in the ranking.

Evidence standard. Every V traces to a trace-baseline delta, not to a benchmark or an analyst assertion. Every R input likewise traces to an observable cost. An error cost that cannot be sourced is recorded as unsourced and surfaced. No risk input is carried to the budget line at its lowest confidence grade.

Outputs. The ρ ledger.

Exit criterion. The workforce partition is complete and every loop's classification carries a recorded, reviewable rationale with the yield risk-adjusted.

PHASE 6 · OPERATING-MODEL ASSEMBLY AND HANDOFF TO TEST

Purpose. Assemble the artifact set and convert the design into testable form.

Activities. Consolidate the six artifacts of 4.7. Instantiate the control catalog (Section 10) against the specific design. Verify postconditions (Section 9). Define the validation-program instance (which hypotheses of Section 11 this deployment will test, with measures and thresholds). Establish the measurement plumbing for the health metrics to set the new baseline.

Exit criterion. Postconditions verified and signed by the executive owner. The design is now an instrument, a set of predictions with measurement attached.

Governance

> *"The agent decided" is not an admissible account of any outcome.*
>
> OOA V1.7, SECTION 6.2

Governance in an agent-native operating model answers one question with three corollaries: who has authority over the policy that governs the agents, and how is that authority exercised, changed, and audited? The model has five components and they exist about the agent and its function. The method by which to build or architect the agent construct is out of scope of this paper.

6.1 Decision rights over the agent layer

The foundational rule, stated in 4.3 and elevated here to a governance principle: agents do not set their own thresholds, and agents do not govern agents. Every parameter that shapes agent behavior is owned by a named human or role in a registered Policy Loop. Ownership means the right to change the parameter, the obligation to review it on the loop's cadence, and accountability for the outcomes the parameter produces.

Parameter changes pass through policy change control. The Policy Loop's periodic cadence is the change-control rhythm. Out-of-cycle changes are permitted but flagged as exceptions and reviewed at the next cycle. The intent is to make operative policy and encoded policy identical by construction, the structural cure for a divergent override rate.

6.2 Accountability architecture

Every loop in the model, agentic or human, has exactly one named accountable human role. For agentic loops, the accountable human role is the Policy Loop owner whose parameters govern it. The agent is an instrument of that person's policy, and "the agent decided" is not an admissible account of any outcome. For human loops, the registry's owner is accountable. For seams, the register's owner is accountable for the friction and for progressing the closure condition.

Three architectural rules prevent diffusion. **Singularity.** One owner per loop. Shared accountability is recorded as no accountability. **Symmetry of information rights.** An owner cannot be accountable for a loop whose state they cannot observe, so every accountability assignment

carries an information-rights specification: what the agents must surface, at what latency. **Non-delegation downward.** The asymmetric boundary of 4.3, agents may hand up and never down, exists so that no human can launder a human judgment through an agent and no agent can launder a commitment through a pro-forma human approval.

6.3 The regulated-edge pattern

Where the seam register identifies regulatory or statutory constraints, the governance pattern changes categorically. For that edge, the continuous-autonomous pattern is replaced by human attestation at the decision point (a named, qualified person commits, with the agent in a preparation role at most, though agentic preparation may be abandoned altogether if R is too high), an immutable ledger for the trail, and hard separation of duties between proposal, commitment, and verification. The design heuristic generalizes: agents handle the unregulated majority, the regulated fraction runs the attestation pattern, and the model is explicit at any given moment about which mode it is in. Ambiguity about mode is itself a control failure (10.2).

6.4 Provenance and audit

Every commitment made by any agent is logged in accordance with enterprise policy and reportable. This decision provenance serves four constituencies: auditors (reconstruction), the Policy Loop (parameter-effect analysis), the Meaning Loop (pattern material), and the validation program (every logged decision is a data point against the Section 11 hypotheses). The same provenance discipline applies to the design artifacts themselves: the set is versioned, its authorship recorded, and its repository access-logged. Operating-model intellectual property is an asset and its provenance is part of the control environment.

6.5 Reversibility

Every agentic loop ships with a reversion path: a documented, rehearsed procedure for returning the loop to human operation, with the staffing and information requirements of reversion specified in advance. Reversibility is a governance requirement for two reasons. First, it is the credible foundation of the kill-switch control (Section 10): a kill switch without a reversion path is a switch to nothing. Second, it disciplines the design: a loop that cannot be specified for reversion is a loop whose human-side knowledge has been allowed to evaporate, which is itself a strategic risk the Policy Loop must explicitly accept or refuse.

Metrics

> *A metric system without Goodhart analysis is an incentive to fail.*
>
> 00A V1.7, SECTION 7

Four families, ordered by the lifecycle: trace metrics describe the current state, collapse metrics state the design's predictions, yield metrics price the allocation, and health metrics monitor the running model. The system closes with named anti-metrics, because a metric system without Goodhart analysis is an incentive to fail.

7.1 Trace metrics (descriptive baseline)

Per loop, from the instrumentation ontology: humans-in-chain; systems-of-record touched; swivel count; stall count and total stall duration; override rate (overridden recommendations divided by recommendations issued); batch-boundary count; consensus-ritual count and hours; elapsed time, trigger to commitment and trigger to fulfillment (median and p90); and the accountability-diffusion indicator (can any single human answer the loop's defining question from system state, yes or no, with time-to-answer if yes).

7.2 Collapse metrics (design predictions)

For each trace metric, the re-derivation states a predicted post-design value with bounds, recorded before deployment. Collapse metrics are the framework's falsifiable forecasts: a deployment whose measured deltas fall outside the stated bounds is evidence against the design or against the trace, and the variance is analyzed. Stating them pre-deployment is the operational content of H2.

7.3 Yield metrics (allocation)

Value (V) is the measured delta against the trace baseline, decomposed into six named mechanisms grouped in three families. **Throughput** carries margin protected, value at risk retired, and stall-time value where time-to-commit has priced consequences. **Investment** carries working capital released, held separately because it is a one-time release rather than an annual flow. **Operating expense** carries labour cost avoided net of seam labour, and error and expedite cost

avoided. The grouping matters because a portfolio whose value is entirely operating-expense reduction is a headcount plan wearing the vocabulary of an operating-model redesign, and reporting V in three blocks makes that visible on the face of the ledger.

The net-of-seam subtraction on labour cost avoided is mandatory. A design that removes human effort from a loop and creates seam review in its place has not avoided that cost, it has moved it, and the gross figure double-counts capacity that came straight back as review. Seam review labour is charged to the lane that creates it.

Cost (C) is fully loaded: inference and orchestration compute, integration build amortized over loop life, run operations, and the human-loop time the agentic loop consumes. Risk (R) is the expected annual cost of the loop's autonomous commitments being wrong, held to the same evidential standard as V and built from the loop's own error structure measured against the trace baseline.

The sourcing discipline is symmetric across all three terms. Each mechanism of V must trace to a baseline observation, and benchmark-derived or asserted value is inadmissible. Each input to R must trace to an observable: a logged override consequence, a true-up delta, a documented escape. A value that cannot be attributed, or an error cost that cannot be sourced, is recorded as such and surfaced, not silently set to zero. Unsourced mechanisms are held out of the ranking, and the ledger reports the unsourced share of V so a reader can see how much of the case rests on value the engagement has not yet evidenced. "Real versus vapor" is the qualification standard for value and risk alike, exactly as for capability. The full computation is Appendix D.

From these, $\rho = V / (C + R)$, the ranking, and the budget line. Because the denominator prices risk, the ranking sorts a loop's value against its full cost of operation including the cost of its errors, and two loops of equal value separate by how dangerous each is when wrong.

Two derived metrics matter at portfolio level. **Yield dispersion** is the ratio of the highest to the lowest ρ across the loops that are funding candidates. High dispersion means the ranking is sorting material differences in value per risk-adjusted cost. Low dispersion means something other than yield chose the portfolio. The measure is taken across candidates rather than across funded loops, because taking it across the funded set alone measures only the loops above the line and cannot detect a budget line that is discriminating between near-identical options. **Constraint shadow price** is the ρ of the best unfunded loop: what one more unit of compute budget would buy, and the Policy Loop's budget argument stated as a number.

7.4 Health metrics and anti-metrics (run-time)

The running model is monitored on six health families.

Boundary integrity. Hand-up rate per agent against design expectation, hand-ups outside declared conditions (a control event), and hand-up SLA attainment by the owning human loop.

Judgment engagement. Per human loop: decision latency, decision distribution (a Trade-off Loop that accepts one hundred percent of agent-presented defaults exhibits a laundering signature), and policy churn in the Policy Loop (zero churn over quarters indicates the loop has stopped governing). The behavioral basis of this measurement is settled across both schools of decision psychology: subjective confidence is produced by fluency and coherence and carries no evidential weight (Kahneman & Klein, 2009), so engagement is measured in decision distributions, never collected as attestation.

Override residual. Post-deployment override rate against the designed exception rate. The anti-metric is explicit: the target is the designed rate, not zero. A zero-override rate is a significant finding because it indicates either perfect policy (rare) or disengaged human judgment (common). See H6.

Drift. Input-distribution shift per agent against validation conditions, decision-distribution shift, and seam-friction trend.

Seam latency. Measured friction cost at each registered seam, trended toward or away from the closure condition.

Provenance completeness. Percentage of commitments with full decision provenance. The target is one hundred.

GOODHART ANALYSIS

Optimizing humans-in-chain to zero destroys the judgment loops (the target is three loops owned). Optimizing hand-up rate downward incentivizes boundary widening without authorization. Optimizing elapsed time alone incentivizes seam denial. Every health metric is paired with its corrupting pressure in the control catalog (Section 10).

Preconditions

> *"Use judgment" is not a policy.*
>
> OOA V1.7, SECTION 8.1

Preconditions are stated in four classes and operationalized as a readiness assessment (Appendix C). The framework's position on unmet preconditions is graduated: some gate the method entirely, others gate only specific phases or convert agentic candidates into seam entries. This graduation is itself part of the method: an unmet precondition is classified and logged.

8.1 Foundation preconditions (per shift, per loop)

For the information shift: the loop's trigger class and information set must be instrumentable, with event streams or polling access to the state that decides the loop. A loop whose information set is fundamentally tacit or relational is not disqualified. It is pre-classified toward human loops by the taxonomy. A loop whose information exists digitally but is access-blocked is a data precondition failure (8.3), which is remediable.

For the coordination and execution shifts: machine-readable interfaces to the systems of commitment. The agent must be able to write the commitment, not draft an email about it. Where the commitment surface is a counterparty's phone tree, that is a seam rather than a precondition failure.

For the cadence shift: the loop's policy π must be expressible as executable rules with quantified exceptions. "Commit unless predicted margin impact exceeds threshold, or SKU is on the watchlist" is executable. "Use judgment" is not a policy, and a loop that cannot state its π beyond "use judgment" routes to Phase 3 for judgment design rather than Phase 2 for agent design.

Instrumentability is necessary and insufficient: the loop's cue-to-outcome structure must also carry learnable regularity (Kahneman & Klein, 2009). A loop whose information is fully digital but whose outcomes resolve too slowly or too noisily for feedback to teach is pre-classified toward human loops, and its human owner's warrant is accountability rather than predictive intuition.

8.2 Organizational preconditions

Condition one is the hard gate: a named executive owner of the operating model as such, with authority to register human loops, assign accountability, and accept the workforce partition. Without this owner, the method produces artifacts without authority and should not proceed. The remainder are willingness to define roles as loops owned rather than tasks performed, since the human resources and role-architecture implication is real and must be sponsored; practitioner access for Phase 1 under the non-audit stance of 7.1; labor-relations and works-council engagement initiated before Phase 5 produces a workforce partition, which jurisdictions with codetermination make a legal precondition; and an agreed value-baseline convention so that Phase 5's V attribution has a settled denominator.

8.3 Data and evidence preconditions

System-log access sufficient to reproduce the trace counts. Timestamp fidelity across systems sufficient to sequence the loop, since clock skew is a real Phase 1 hazard. Access to the exception path, not only the curated happy path. And retention adequate to observe at least several full loop cycles. Where logs cannot support a count, the count is taken by direct observation and labeled as such. Interview-sourced counts are inadmissible (Phase 1 evidence standard).

A published business capability model with a stated version and an owning architecture authority is a further precondition of this class, and it classifies rather than gates. Where one exists, the capability anchor is recorded against each loop in the inventory and the design can be placed in the enterprise's portfolio at handoff. Where none exists, or where the published model is stale enough that its owning authority will not stand behind it, the absence is recorded as absent and the engagement proceeds. An anchor that cannot be sourced is never invented, and the ordering condition means no engagement waits on one.

8.4 Economic preconditions

A compute budget that is bounded, since the budget-line construct degenerates if compute is treated as free, and a costing model capable of loading C accurately. An organization that cannot price its compute cannot run the ρ allocation and will, predictably, fund agentic loops by enthusiasm, which is the platform's allocation method reimported. This is the economic form of the ordering condition (1.2): without a bounded budget and an accurate cost model, the negative conclusions the method depends on cannot be reached numerically.

Postconditions

> *No agent exists for which "what will it not decide?" lacks a written answer.*
>
> OOA V1.7, SECTION 9, P1

The verifiable state of the design at the exit of Phase 6, stated against the artifact set. They are written as assertions. Each must be checked, and the executive owner signs the set.

TABLE 9 · POSTCONDITIONS P1 TO P8

ID	POSTCONDITION	ASSERTION
P1	Boundary completeness	Every agent has a complete judgment boundary: scope, explicit non-scope, typed hand-up conditions with SLAs, and a named receiving human loop. No agent exists for which "what will it not decide?" lacks a written answer.
P2	Judgment closure	Every hand-up condition maps to exactly one registered human loop. Every loop has a singular named owner, a cadence, decision rights, and information rights. The three-loop taxonomy classifies every loop, with any residuals documented as taxonomy counterexamples.
P3	Seam closure	The derived flow plus the seam register account for the loop's full path. Every seam has a pattern, an owner, a measured friction, and a closure condition. Regulated seams instantiate the attestation pattern in full.
P4	Allocation closure	Every loop in the inventory is classified, agentic, human, or eliminated, with a recorded rationale. The ρ ledger shows the ranking, the budget line, the constraint flag, and the shadow price. No loop is funded by assertion.
P5	Control instantiation	The control catalog of Section 10 is instantiated against this design. Every control has an owner, a mechanism, and a test. The kill switch and reversion path for every agentic loop are documented and rehearsed at least once before go-live.
P6	Measurement readiness	The health-metric plumbing is live before the loops are. Collapse-metric predictions are recorded with bounds and timestamps pre-deployment. Decision provenance is capturing at one hundred percent in pre-production.
P7	Reversibility	Every agentic loop's reversion path is specified, staffed on paper, and rehearsed. The knowledge-evaporation risk of each loop has been explicitly accepted or mitigated by the Policy Loop owner.

ID	POSTCONDITION	ASSERTION
P8	Provenance of the design	The artifact set is versioned, authored, repository-controlled, and access-logged from first draft. Where a published capability model exists, every loop in the inventory carries its capability anchor with the owning authority and the model version.

Controls

> *"No change, considered" is a record. "No meeting" is a finding.*
>
> OOA V1.7, SECTION 10.3

Three tiers, design-time, run-time, and periodic, following the model-risk grammar. Every control names its risk, its mechanism, and its test.

10.1 Design-time controls

The agentic-line test RISK: UNBOUNDED AGENCY

No agent passes design review without a complete judgment boundary.

Test. An independent reviewer attempts to state, from the artifact alone, what the agent will not decide. Failure to state it is failure of the review.

Forcing-reason audit RISK: SHAPE INHERITANCE, THE CONWAY TRAP

Every structure in the derived flow carries a recorded force, either the loop or a declared seam.

Test. Sample structures. Any structure whose recorded force is 'current practice' fails.

Inverse test RISK: METHOD BIAS TOWARD AUTOMATION, ORDERING-CONDITION FAILURE

At least one loop in every engagement's inventory is run through the full method where the expected outcome is human ownership or elimination. The method recommending agents for every loop it touches is treated as evidence of analyst capture, and the ρ ledger is re-reviewed.

Test. Presence and disposition of the inverse case. An engagement with no human-owned or eliminated outcome is a finding against the ordering condition.

Anchor direction test RISK: TAXONOMY CAPTURE

Where a capability anchor is recorded, the reference runs from the loop out to the capability model and never the reverse. A loop enters the inventory because a commitment recurs, and never because a capability-model node exists to be covered. The control screens the direction of the reference. It does not question whether the enterprise should be anchored to at all.

Test. Sample loops from the inventory and confirm each traces to a commitment. A loop whose recorded origin is a capability-model entry fails.

Decision laundering pre-test RISK: PRO-FORMA JUDGMENT

Every human loop's specification includes the observable difference between engaged and rubber-stamped exercise, and the health metric that would detect it.

Test. Specification review.

Separation of design duties RISK: SELF-CERTIFICATION

The analyst who derives the flow does not solely verify the postconditions. P-series verification requires a second reviewer.

Test. Reviewer-of-record check on the signed postcondition set.

10.2 Run-time controls

Boundary enforcement RISK: SCOPE CREEP IN PRODUCTION

Agent commitments are checked against the declared boundary at execution time and antagonized by both policy and technology guardrails. An out-of-boundary commitment attempt is blocked, logged, and paged to the Policy Loop owner.

Test. Injected boundary-violation attempts in pre-production. Zero pass-throughs.

Hand-up integrity RISK: SILENT ABSORPTION OF EXCEPTIONS

Hand-ups outside declared conditions, and declared conditions that fire without a hand-up, are both control events with automatic escalation. The control governs exception and trade-off hand-ups. Signal hand-ups (4.3) are exempt by construction and are reconciled instead for completeness of record.

Test. Log reconciliation of condition evaluations against pages, and completeness reconciliation of the signal channel.

Drift detection RISK: SILENT ENVIRONMENT SHIFT

Input and decision distributions are monitored against validation baselines per agent. Breach of drift thresholds suspends autonomous commitment for the affected agent and routes its proposals through the reversion path until revalidated.

Test. Synthetic drift injection.

Kill switch and circuit breakers RISK: CORRELATED OR RUNAWAY FAILURE

Per-loop kill switch exercising the rehearsed reversion path. Portfolio-level circuit breakers on aggregate commitment value per unit time, with thresholds owned by the Policy Loop.

Test. Periodic kill-switch drills, timed.

Provenance capture RISK: UNRECONSTRUCTABLE DECISIONS

A commitment is blocked if the provenance write fails. Provenance is in the transaction.

Test. Provenance-store fault injection. Commitments must fail closed.

Regulated-mode assertion RISK: MODE AMBIGUITY AT REGULATED EDGES

Every commitment in a loop touching a regulated seam carries an explicit mode flag, autonomous or attestation. Attestation-mode commitments without a named attestor fail closed.

Test. Edge-case injection at the seam.

Policy version binding RISK: UNTRACEABLE PARAMETER EFFECTS

Every commitment binds the policy version. Policy changes propagate atomically, so no agent runs a mixture of versions within a commitment.

Test. Change-propagation audit during a controlled parameter change.

10.3 Periodic controls

Policy Loop cadence enforcement RISK: GOVERNANCE ATROPHY

The Policy Loop's review occurs on cadence and its minutes record parameters reviewed, changed, and deliberately left unchanged. “No change, considered” is a record. “No meeting” is a finding.

Test. Cadence-adherence and minute-completeness review.

Revalidation RISK: STALE VALIDATION

Each agent is revalidated on a fixed clock and on any material policy or environment change.

Test. Revalidation-record audit against the clock.

Seam reassessment RISK: OSSIFIED ACCOMMODATION

Each seam's closure condition is tested on cadence. A seam whose counterparty has shifted but whose accommodation persists is a finding.

Test. Closure-condition test per seam.

Laundering audit RISK: JUDGMENT DECAY

The engagement-defined laundering signatures of 7.4 are reviewed per human loop, with owner rotation or information-rights redesign as the corrective.

Test. Decision-distribution review (see 11.2).

ρ re-ranking RISK: STALE ALLOCATION

The ledger is re-ranked on cadence with realized V replacing predicted. Loops whose realized yield falls below the shadow price are candidates for defunding.

Test. Re-rank reconciliation against realized value.

Validation and testability

> *The framework is validated the way it works: by trace, prediction, and registers.*
>
> OOA V1.7, SECTION 11.3

The framework's claims are stated here as hypotheses with measures, so that the conversion from framework to instrument is a matter of operationalization rather than reinterpretation. The validation program has two layers: method validity (is OOA measuring and predicting what it claims, anywhere?) and deployment validity (did this design do what it predicted, here?). Each engagement contributes data to both.

11.1 Hypotheses

H1 · PORTABILITY (SECTOR-AGNOSTICISM)

The method applies to a decision loop in a second, unrelated domain with zero new primitives: every construct needed is already in the ontology, and no domain forces an extension.

MEASURE	Primitive-extension count across domains.
THRESHOLD	Zero forced extensions across at least three domains.
FALSIFIER	Any domain whose loops cannot be stated as five-tuples, or whose trace requires new counting constructs.

> The measurement period for H1 has not opened. Step (i) of the path from framework to instrument (11.3) is to freeze the ontology and the counting rules of 4.1 as a coding manual, and that step remains future work. Vocabulary added before the freeze is recorded in the extension log below and is not counted against H1. H1 begins measuring at the freeze.

EXTENSION LOG · VOCABULARY ADDED BEFORE THE ONTOLOGY FREEZE

TERM	RECORDED	KIND	REASON
Work capability	2026-07-30	Primitive	Cross-artifact reconciliation. A structural layer beneath the decision loop (3.2) and above the agent specification, which both instruments had invented independently under different names.
Capability anchor	2026-07-30	Reference field	Cross-artifact reconciliation. A reference out to a document the method does not own, carrying no design authority, closer in kind to the system-of-record touch construct already in the 4.1 ontology than to a building block.

> Neither entry arose from applying the method in a new domain, which is what H1 measures. Both arose from testing the paper against its own instruments. The log exists so that the freeze, when it comes, can be audited against a record of what moved beforehand.

H2 · PREDICTIVE VALIDITY OF THE COLLAPSE

Pre-deployment collapse predictions (7.2) fall within stated bounds post-deployment.

MEASURE	Per-metric prediction error across deployments.
THRESHOLD	Engagement-defined bounds, recorded before go-live.
FALSIFIER	Systematic directional bias, where the method consistently over-promises a given metric, indicating a defect in the re-derivation discipline.

H3 · EXHAUSTIVENESS OF THE HUMAN-LOOP TAXONOMY

All hand-up events and all designed human judgments classify into Policy, Trade-off, or Meaning loops.

MEASURE	Residual, meaning unclassifiable, rate across the hand-up corpus of running deployments.
THRESHOLD	Residual below a small engagement-defined epsilon, with every residual documented as a counterexample.
FALSIFIER	A recurring, coherent class of human judgment that resists all three categories. This is the most theoretically valuable negative result the program can produce.

H4 · DISCRIMINANT VALIDITY (ORDERING CONDITION)

The method recommends against agentic deployment where the constructs say it should: low-ρ loops, non-optimizable decision variables, regulated edges. That a loop can be agentified does not establish that it should be.

MEASURE	Distribution of loop classifications across engagements.
THRESHOLD	A method whose output distribution approaches “agentic everywhere” fails discriminant validity regardless of client satisfaction.
FALSIFIER	A consistent absence of human-owned or eliminated outcomes across engagements, the empirical signature of an inverted ordering condition.

H5 · RELIABILITY OF THE TRACE

Independent analysts applying the instrumentation ontology to the same evidence produce the same counts.

MEASURE	Inter-rater agreement on the trace constructs.
THRESHOLD	Agreement high enough that trace metrics are reportable as measurements, with residual disagreement amounting to classification noise.
FALSIFIER	Persistent disagreement on a construct, indicating the counting rule is underspecified and returning the construct to definition.

H6 · OVERRIDE CONVERGENCE

Post-deployment override and exception rates converge toward the designed rate rather than toward zero, and divergence in either direction predicts a diagnosable cause.

MEASURE	Override-residual trajectory and its correlation with Policy Loop churn and laundering-signature metrics.
THRESHOLD	Convergence toward the designed rate, with upward divergence predicting a policy-reality gap and downward-to-zero divergence predicting laundering.
FALSIFIER	Override residual uncorrelated with any diagnosable cause, which would weaken the claim that the designed exception rate is a meaningful target.

11.2 Construct-validity notes

Three constructs carry known measurement hazards, recorded here so the instrument design confronts them rather than inherits them.

V attribution is the hardest. Value deltas against a baseline are confounded by everything else the enterprise does simultaneously. The instrument will need holdout loops, phased rollouts with comparison windows, or explicit confound registers, and the discipline that an unattributable V is recorded as unattributable.

Stall measurement depends on distinguishing "information existed but waited" from "information did not exist". The counting rule requires evidence of existence and not an assumption of it.

Laundering signatures risk insulting the humans they monitor. The instrument measures decision distributions, and the 7.1 non-audit stance applies to run-time judgment metrics as forcefully as to the trace.

11.3 Instrumenting the framework

The unpacking path this paper is designed to feed: (i) freeze the ontology and counting rules of 4.1 as a coding manual; (ii) build the trace as a structured data capture (the instrumentation schema of Appendix B); (iii) pre-register collapse predictions per engagement; (iv) accumulate the cross-engagement corpus on which H1 to H6 are tested; (v) treat every taxonomy residual, prediction miss, and forced extension as program data. The framework is validated the way it works: by trace, prediction, and registers.

Failure modes and limitations

> *It fails while reporting success.*
>
> OOA V1.7, SECTION 12.1, ON JUDGMENT LAUNDERING

12.1 Failure modes

Each failure mode has an associated control. The authors acknowledge that this is neither exhaustive nor complete, and that real-world scenarios will surface additional modes to be documented and controlled.

TABLE 12.1 · FAILURE MODES AND THEIR CONTROLS

FAILURE MODE	DESCRIPTION	CONTROL
Conway trap	Agents deployed along the current organization chart replicate siloed work as siloed automation. This is the failure mode of the retrofit. Within an OOA engagement it reappears whenever Phase 2's subtractive discipline slips.	Forcing-reason audit (10.1).
Judgment laundering	The human loops decay into rubber stamps and the accountability architecture becomes ceremonial while the risk profile becomes that of full autonomy without the design for it. This is the failure mode the framework fears most, because it fails while reporting success. The mechanism is documented rather than hypothetical: automated support systems produce passivity in the humans who oversee them, which is the condition rubber-stamping requires (Klein, Snowden, & Chew, 2007).	Laundering pre-test (10.1), engagement metrics (7.4), periodic audit (10.3).
Seam denial	The design pretends a counterparty has shifted. The integration friction can make shifting a business unit undesirable and potentially not feasible.	Evidence-based seam classification (Phase 4). The qualification register surfaces the risk early.
Accountability diffusion	The old chain's diffusion re-emerges as "the agent decided".	Singularity of ownership, non-delegation downward, provenance (6.2, 6.4).
Metric gaming	Optimizing humans-to-zero, hand-ups-to-zero, or elapsed-time-at-any-cost.	Each paired with its anti-metric (7.4).

FAILURE MODE	DESCRIPTION	CONTROL
Policy ossification	The encoded policy outlives the world it was written for, and because overrides are now governed, the divergence accumulates until it surfaces as drift or as a Trade-off Loop drowning in hand-ups.	Policy Loop cadence enforcement (10.3). A governance body that has stopped asking its defining question has stopped existing.
Knowledge evaporation	The human capacity to operate the loop manually atrophies and reversion paths become fictional.	Rehearsed reversion, explicit acceptance of residual risk by the owner (6.5).
Taxonomy capture	The decision inventory is populated from an inherited enterprise taxonomy rather than from observed commitments. The capability model, the process reference model, or the organization chart supplies the list of things to consider, and the loops that exist in the work but not in the taxonomy are never seen. The engagement then produces a design that is conformant to the enterprise's existing map and blind in the same places the map is blind. This is the ordering condition of 1.2 failing through an artifact that arrives with architectural authority rather than through platform economics.	Anchor direction test (10.1).

12.2 Limitations

As with any framework, limitations exist and the model will not apply in all scenarios.

Partial foundations bound the yield. The shifts are partial and uneven, and the method manages, rather than abolishes, the resulting friction.

Value attribution is genuinely hard. The insistence on baseline-traceable V will sometimes mean reporting less value than is real, because the real value is unattributable. This is a feature with a cost.

The decision lens underweights tacit coordination. Not all meetings should be abolished. Meetings that act as trust maintenance and sense-making have inherent value that feeds the organization's knowledge culture. A1's decomposition is, at the margins, an idealization. The Meaning Loop absorbs some of this, but culture is not in the ontology.

Regulated edges cap the design space by construction, and jurisdictional variance (codetermination, sectoral regulation, data residency) means the same loop re-derives differently across a multi-country footprint, handled through per-jurisdiction seam entries, at true analytical cost.

The evidence base is young. The cross-engagement corpus is at the beginning of its accumulation. A synthetic corpus verifies that the economic model's machinery behaves correctly under varied inputs, and it is not evidence about business models and must never be described as such.

Conclusion

> *When the foundation of work changes, re-derive from the decision up.*
>
> OOA V1.7, SECTION 13

The framework makes one central methodological claim and builds a procedure around it: when the execution foundation of work changes, the method re-derives operating models from the decision up rather than retrofitting them from the org chart down, and the re-derivation is incomplete until human judgment is designed, seams are recorded, compute is allocated by yield, and the whole is placed under governance, measurement, and control rigorous enough to be falsified.

Its neutrality is structural and conditional: the decision precedes the platform, and precedes every other inherited enterprise taxonomy; negative conclusions are admissible; the qualification registers are mandatory; and where the ordering condition is not satisfied, the artifacts reveal it. Its sector-agnosticism is definitional and testable: the primitives quantify over no domain. Its accountability is architectural: every loop has one named human, agents never set their own thresholds, and "the agent decided" is never an account of anything. And its epistemic posture is the one it imposes on its deployments: predictions registered before outcomes, residuals recorded rather than narrated, and a standing invitation to be proven wrong in the specific places, H1 through H6, where being wrong would teach the most.

What this paper deliberately does not do is operationalize. The coding manual, the capture schema, the pre-registration templates, and the cross-engagement corpus are the next movement, the conversion of the framework into the testable instrument it is written to become. When the foundation of work changes, re-derive from the decision up. Design the human judgment. Record the seams. Allocate by yield. Then place the whole under governance rigorous enough to be falsified.

Glossary (locked lexicon)

> *The decision precedes the platform. Negative conclusions are admissible.*
>
> OOA V1.7, CLOSING STATEMENT

Outcome-Owned Architecture (OOA) the noun; the standing operating model this paper articulates, an enterprise re-derived from the decision up. The masthead construct, of which the design act and the economic model are parts.

Outcome-Earned Economic Model the economic engine of OOA; the compute-yield allocation ρ = V / (C + R) that ranks decision loops by value per unit of risk-adjusted cost and derives the workforce from the budget line.

Execution foundation (locked term; supersedes *substrate* from earlier trace drafts) the assumed physics of work along F1 to F4.

Decision loop recurring commitment of resources under uncertainty; five-tuple ⟨`trigger, information set, policy, commitment, feedback`⟩.

Foundation shift qualitative change in a foundation dimension (instantaneous information; unimpeded coordination; autonomous execution; continuous decisions).

Trace instrumented current-state reconstruction of one loop's lifecycle.

Swivel manual transfer of loop state between systems.

Stall waiting on batch, calendar, or human while the required information already exists.

Override human reversal of a system recommendation; rate measures encoded-versus-operative policy divergence.

Judgment boundary per-agent declaration of non-scope and typed hand-up conditions.

Human-loop taxonomy Policy Loop (rules governing agents), Trade-off Loop (non-optimizable commitments), Meaning Loop (pattern interpretation); claimed exhaustive (H3).

Seam interface with an unshifted foundation; the seam register is its governed inventory with patterns, owners, and closure conditions.

Regulated-edge pattern attestation at the decision point, immutable ledger, separation of duties.

Compute yield (ρ = V / (C + R)) the Outcome-Earned Economic Model's core ratio: value per unit of risk-adjusted compute; the budget line is the funding cutoff from which workforce definition is derived.

Judgment laundering decay of human loops into rubber stamps.

Disposition the assignment of a single decision loop to agent ownership or to human ownership, together with the warrant that justifies it. Assigned at the loop (3.2) and never at the role, because expertise fractionates by task. The disposition taxonomy is the full set of these assignments across the decision inventory.

Work capability bounded work that creates, updates, validates, or explains the state a decision loop's information set requires, or that executes a commitment the loop has made. Named as verb plus object plus output. Role-neutral and system-neutral. Derived from the loop and never inherited from the organization chart. Sits beneath the decision loop in the artifact hierarchy and above the agent specification.

Capability anchor the reference from a decision loop, or from a work capability beneath it, to a leaf node in the enterprise's own published business capability model at whatever level that model publishes, carried together with the owning architecture authority and the model version. The anchor is a mapping output of the method. It carries no design authority.

Real versus vapor evidence standard for capability and value claims alike.

Collapse metrics pre-registered design predictions against trace baselines.

Ordering condition the requirement that the decision inventory precede platform selection and every other inherited enterprise taxonomy; the structural basis of OOA's neutrality (1.2, A6).

Agentic Workforce Design (the design act) the design act that produces an OOA. Held as internal working vocabulary for the discipline that stands up an OOA. OOA is the front-facing noun and carries the framework throughout this paper.

Instrumentation schema (Phase 1 capture)

Per loop instance, the trace captures, as structured records rather than narrative:

- **Loop header.** Loop id, five-tuple, domain, instance class (median, p90, exception), and the capability anchor where one exists.
- **Step records.** Timestamp, actor type and role, action class, systems touched, artifacts produced, evidence source (log, observation, interview-interpretation).
- **Swivel records.** From-system, to-system, state transferred, duration.
- **Stall records.** Start and end, foundation dimension compensated, upstream-existence evidence for the waited-on information.

- **Override records.** Recommendation, action taken, stated reason class, downstream consequence if observable.
- **Ritual records.** Gathering, duration, attendance, output state.
- **Terminal record.** Defining question, who was asked, time-to-answer or failure.

Counting-rule resolutions adopted during the engagement are appended to the schema and version-controlled with it, so that H5 reliability testing runs against a fixed manual.

Readiness checklist

GATE ITEMS (ENGAGEMENT DOES NOT PROCEED WITHOUT)

- Named executive owner of the operating model with loop-registration authority.
- Phase 1 evidence access (logs, observation, exception path).
- Agreed value-baseline convention.
- Front-of-work gate executed (role, work-breakdown ownership, IP-provenance conventions confirmed in writing).

PER-LOOP ITEMS (CLASSIFY RATHER THAN GATE)

- F1, trigger and information set instrumentable (else: pre-route to judgment design).
- F2 and F3, commitment surface machine-writable (else: seam entry).
- F4, policy statable as executable rules with quantified exceptions (else: Phase 3 routing).
- Regulated constraints identified (else-risk: mode ambiguity, resolve before Phase 2).
- Capability anchor recorded, with the owning architecture authority and the model version (else: recorded as absent).

ORGANIZATIONAL ITEMS (SCHEDULE-CRITICAL)

- Role-architecture sponsorship for loops-owned definitions.
- Labor-relations and works-council engagement initiated before Phase 5 in applicable jurisdictions.
- Practitioner access secured under the non-audit stance.
- Compute budget bounded and costing model able to load C.
- Published business capability model with a stated version and an owning authority (classify rather than gate).

The compute-yield allocation

> *A loop whose haircut is implausibly low has usually had a severity or an error rate set by optimism.*
>
> OOA V1.7, APPENDIX D.5

The method's mechanism for deriving the workforce from the economics of the loops rather than from the org chart. It ranks candidate loops by ρ, funds them in descending order against a bounded compute budget, and reads the workforce partition off where the funding line lands. This appendix specifies each term and the ranking at the level of detail the body does not carry, because this is the part of the method most exposed to scrutiny and a reader should be able to audit it.

D.1 Value, V

V is the value attributable to agentic operation of the loop, measured as a delta against the trace baseline (7.3), decomposed into six named mechanisms in three families. Reporting the families separately is load-bearing: a portfolio whose value is entirely operating-expense reduction is a headcount plan in the vocabulary of an operating-model redesign, and the split makes that legible on the face of the ledger.

TABLE D.1 · THE SIX VALUE MECHANISMS

FAMILY	MECHANISM	DEFINITION AGAINST THE TRACE BASELINE
Throughput	Margin protected	Margin retained on commitments that would have leaked under the baseline policy, measured per commitment against the trace's realized leak. Benchmark-derived margin is inadmissible.
	Value at risk retired	Exposure removed because a commitment is now made on current state rather than stale state. Traces to a baseline exposure record, and is recorded as unattributable where no record exists.
	Stall-time value	Value of commitments reached earlier, where time to commit carries a priced consequence. Where delay carries no price the mechanism is zero, and is recorded as zero rather than omitted.

FAMILY	MECHANISM	DEFINITION AGAINST THE TRACE BASELINE
Investment	Working capital released	Cash freed by shorter commitment cycles or reduced buffer. Traces to a balance-sheet observation. Held separately from throughput because it is a one-time release rather than an annual flow.
Operating expense	Labour cost avoided, net of seam labour	Loaded human cost of the work the agent takes, less the loaded cost of the seam review the redesign creates. Volume, minutes per task, and loaded rate all sourced. The net-of-seam subtraction is mandatory, because the gross figure double-counts capacity that came straight back as review, and the seam labour is charged to the lane that creates it.
	Error and expedite cost avoided	Reduction in rework, expedite premiums, and error remediation against the baseline error rate. Traces to a logged error consequence or an expedite invoice.

Each mechanism traces to a baseline observation. The discipline is that an unattributable value is recorded as unattributable rather than allocated by narrative. Unsourced mechanisms are held out of the ranking, and the ledger reports the unsourced share of V at portfolio level. The method will sometimes report less value than is real because the real value cannot be sourced, and it accepts that cost (12.2).

D.2 Cost, C

C is the fully loaded cost of running the loop agentically: inference and orchestration compute, integration build amortized over loop life, run operations, and the human-loop time the agentic loop consumes. Policy, Trade-off, and Meaning attention is a cost of the loop, not free. C is cost to run. It does not include the cost of the loop being wrong, which is carried separately as R so that the two are visible and sourced independently.

D.3 Risk, R

R is the expected annual cost of the loop's autonomous commitments being wrong. It is not a flat haircut. It is built per loop from the loop's own error structure, along two axes the design holds separate, and is sourced from observables against the trace baseline under the same evidence discipline as V.

Control, how the commitment was governed. A rule-bounded commitment executes deterministic policy inside a guardrail; when it errs, it errs inside the policy band, and its cost is the in-band error net of what is recovered at the next true-up. A true-judgment commitment is an unsupervised genuine call; when it errs, it carries the full cost of the misjudged commitment. The rule-bounded share of a loop's autonomous slice is an evidenced quantity, not an assumption.

Nature, how the error reaches the world. A transactional commitment is one the agent executes directly, and its error realizes directly as a margin or value leak. An informational commitment is one where the agent informs a human who then acts, and its error realizes only through the chain that follows. It is priced as exposure × P(the error changes the human action) × (1 – recovery) × realized-loss fraction.

The three error streams. An agent that only informs cannot, by itself, cause the loss. The human action and the recovery opportunities between the agent and the outcome attenuate it. For each loop, R sums three streams, each a frequency times a severity: rule-bounded autonomous error (autonomous tasks × rule-bounded share × autonomous error rate, at in-band severity); true-judgment autonomous error (autonomous tasks × judgment share × autonomous error rate, at true-judgment severity); and seam escape (ratified tasks × residual escape rate, at true-judgment severity), the commitments that pass human review and are wrong anyway. The seam pattern (4.5) and the regulated edge (6.3) reduce the third stream to its residual but do not eliminate it. R is the sum of the three.

Two invariants the structure asserts, both of which an implementation must enforce and a reader can check.

1. **Attenuation.** A transactional lane takes an action-change probability of one and a recovery of zero. A transactional commitment realizes its error directly, so neither the human-action term nor the recovery chain attenuates it, and any calibration that applies attenuation to a transactional lane has mis-specified the lane.
2. **Severity ordering.** True-judgment severity exceeds in-band severity on every lane. An unsupervised judgment error carries the full cost of the misjudged commitment while a rule-bounded error carries only the in-band cost net of true-up, so a calibration that reverses the ordering has inverted the two.

Why the structure makes R a control. A loop reduces its R by narrowing what it commits autonomously (raising the ratified share), by tightening the policy band (raising the rule-bounded share), or by lowering the severity of being wrong (reversibility, true-up recovery, separation of duties), the same levers the judgment boundary (4.3) and seam register (4.5) already pull, now priced into the loop's yield.

Severity is sourced, not assumed. Each severity traces to an observable: in-band severity to a true-up log; true-judgment severity to the realizable margin leak for transactional loops, or to the exposure-and-recovery chain for informational ones. The two load-bearing assumptions, the

rule-bounded share and the realized-loss fraction, are stated explicitly and sourced rather than buried in a blended rate. Every risk input carries a confidence grade (sampled from logs, benchmark, or estimate), and no risk input is taken to a client at the lowest grade.

D.4 The ranking, the budget line, and the constraint

Loops are ranked by ρ descending. The judgment constraints of A4 are applied first: a loop assigned to humans by the taxonomy is not a funding candidate regardless of its ρ. The remaining loops are funded in descending ρ order until the compute budget is exhausted. The position where funding stops is the budget line, and it derives the workforce: loops above the line are agentic, loops below it remain human or are eliminated where the trace shows they were pure foundation compensation.

Two portfolio metrics read the result. **Yield dispersion** is the ratio of the highest to the lowest ρ across funding candidates. High dispersion means the budget line is doing real work. **Constraint shadow price** is the ρ of the best unfunded loop: at the budget line, it is the yield the enterprise forgoes on the best loop it chose not to fund, and therefore the statement of what one more unit of compute budget is worth. The shadow price is the number the Policy Loop carries into the budget argument.

The ranking does not locate the constraint, and the ledger must not pretend otherwise. ρ measures value per unit of risk-adjusted cost. It does not measure whether a loop sits at the queue that governs throughput. A high-yield loop away from the binding constraint can be funded first and buy no throughput at all, which is the Theory of Constraints result stated in Section 2 applied to the method's own instrument. The ledger therefore carries a constraint flag beside each rank, recording whether the loop sits at the binding constraint, with the throughput and operating-expense split of its V beside it.

Where the top-ranked loop relieves no constraint and the loop that does falls below the budget line, the ledger records the conflict and hands it to the Policy Loop owner. It does not resolve it arithmetically. Two honest readings exist in that situation, that the constraint lane should be funded ahead of its rank, or that the constraint should be relieved by means other than compute, and choosing between them is a Trade-off Loop decision by construction (4.4). An instrument that silently reordered the ranking to favour the constraint lane would be making that call on the enterprise's behalf and hiding it.

D.5 Counter-checking, and the canonical figures

The ratio R / V per loop is the implied risk haircut, the fraction of a loop's gross value consumed by the expected cost of its errors. Reading it against intuition, where a high-volume rule-bounded loop should haircut in the low single digits and a high-severity judgment loop should haircut materially more, is the first check that the risk build is calibrated rather than mechanical. A loop whose haircut is implausibly low has usually had a severity or an error rate set by optimism. A loop whose haircut approaches or exceeds its value is a loop the ranking will correctly refuse to fund, and the refusal is the term doing its job.

CANONICAL FIGURES · SCHEMA-CONFORMANT · FROZEN REFERENCE SCENARIO

$24.2M	$36.2M	2.35yr	3.74×	$17.7M
RISK-LOADED NET PRESENT VALUE	CLEAN NET PRESENT VALUE	PAYBACK PERIOD	PORTFOLIO YIELD ρ SOURCED VALUE, ALL LANES	TENTH PERCENTILE STD. ERROR $429,317

Stamped scenario: OEEM v5.1 · entropy frozen 2026-07-30 · distribution method at D.6

The portfolio yield figure is quoted on sourced value across all lanes. Both qualifiers are load-bearing. Sourced value excludes mechanisms the engagement has not yet evidenced, and the all-lanes form spans every lane in the ledger rather than only those passing the candidacy filter. The two forms coincide in this scenario because every lane is a candidate, and they separate in any engagement where the judgment constraint of A4 removes a lane from candidacy. A yield figure quoted without both qualifiers is ambiguous.

D.6 The distribution behind the tenth percentile

The tenth percentile is the one canonical figure whose job is to carry uncertainty, so the method behind it is stated rather than assumed.

TABLE D.6 · SCENARIO STAMP

SAMPLER	Inverse-transform, one uniform per draw per driver, read by both the side test and the placement term
GENERATOR	numpy PCG64
SEED	20260730

TRIALS	1,000
ENTROPY FROZEN	2026-07-30, checksum 3010.298924
TENTH PERCENTILE	$17,687,194
STANDARD ERROR OF P10	$429,317
DISTRIBUTION STANDARD DEVIATION	$8,508,306
P(NPV > 0)	0.999

The entropy is frozen rather than the outputs. Freezing the draws makes the distribution reproducible while leaving it responsive to changes in the driver ranges and the model anchors, which is the property a scenario stamp needs. Freezing the results instead would produce a figure that no longer moves when the model does, which is worse than having no stamp at all.

Two integrity checks guard the stamp: one confirms the live tenth percentile reproduces the stamped value, and one confirms the frozen entropy block is intact by checksum. A failure of either means the draws have moved and any figure quoted from that file is stale.

One methodological commitment travels with this appendix. Draws are never re-run until a sample matches a published number. The gap between an earlier published $19.2M and the drawn $18.83M was 0.78 standard errors, which means selecting a favourable draw would have succeeded on the first or second attempt and produced a figure that looked identical to an honest one. The protection against that is procedural rather than statistical: the entropy is stamped, the checksum is published, and a figure that does not reproduce from the stamp is restated rather than redrawn.

References and intellectual lineage

> *Complexity frequently takes the form of hierarchy.*
>
> SIMON, "THE ARCHITECTURE OF COMPLEXITY," 1962

A companion note to Section 2. It gives the full citation for each source OOA invokes and explains what the method borrows, adapts, adds, or rejects. The list is limited to the works the paper relies on, especially where OOA extends older ideas about work, systems, human judgment, governance, and control into an agentic setting. It is not a survey of the current agentic-AI literature.

THEORY OF CONSTRAINTS AND THROUGHPUT ACCOUNTING

Goldratt, E. M., & Cox, J. (1984). *The Goal: A Process of Ongoing Improvement.* North River Press. ISBN 978-0-88427-178-9.

Supplies the grounding move: the binding constraint is the object of management, the five focusing steps are the procedure for managing it, and improving anything other than the constraint is local optimization. The method generalizes the constraint from the factory floor to the decision queue, names the seam as the post-re-derivation constraint, and reads the budget line and shadow price as constraint economics on the decision portfolio. It also takes the warning: a ranking that funds a non-constraint buys no throughput, which is why D.4 carries the constraint flag.

Goldratt, E. M. (1997). *Critical Chain.* North River Press. ISBN 978-0-88427-153-6.

Extends the grounding in two directions the method relies on. It locates the binding constraints of mature organizations in policy rather than physical capacity, so a constraint does not dissolve when capacity arrives and must be identified and deliberately removed, with the standing warning that inertia converts yesterday's solution into today's constraint. And it supplies the mechanism by which inherited structure accumulates its bulk: safety embeds at every level, each layer adding its own buffer, review, and margin, so the structure grows by locally rational increments that no local actor has the authority to remove. That is the shape the re-derivation subtracts, and the reason A3 expects retrofit to preserve it.

> *Author note. W. H. Evans is a certified Jonah in the Theory of Constraints, Goldratt Institute. The grounding of the method in constraint theory is drawn from that practice.*

PROCESS REDESIGN AND THE RE-DERIVATION

Hammer, M. (1990). "Reengineering Work: Don't Automate, Obliterate." *Harvard Business Review*, 68(4).
Hammer, M., & Champy, J. (1993). *Reengineering the Corporation.* HarperBusiness.

The closest precedent for OOA's re-derivation move. OOA departs in three respects: it takes the decision loop rather than the process as its unit of analysis; it treats human judgment loops as designed first-class outputs rather than cost to be engineered out; and it replaces un-instrumented redesign with a measured trace and bounded, pre-registered predictions.

LEAN, VALUE-STREAM MAPPING, AND FLOW EFFICIENCY

Rother, M., & Shook, J. (1999). *Learning to See.* Lean Enterprise Institute. ISBN 978-0-9667843-0-5.
Womack, J. P., & Jones, D. T. (1996). *Lean Thinking.* Simon & Schuster. ISBN 978-0-7432-4927-0.

Contribute the discipline of walking the actual flow and naming waste. The trace's swivel and stall constructs are recognizable relatives of motion- and waiting-waste.

Reinertsen, D. G. (2009). *The Principles of Product Development Flow.* Celeritas. ISBN 978-1-935401-00-1.
Modig, N., & Åhlström, P. (2012). *This Is Lean.* Rheologica. ISBN 978-91-980393-4-3.

Together these supply the queueing economics and the resource-versus-flow-efficiency distinction behind the trace's wait claim: in knowledge work, waiting dominates lead time, and the value-adding touch is a small share of the whole.

TECHNOLOGICAL REVOLUTIONS AND TIMING

Perez, C. (2002). *Technological Revolutions and Financial Capital.* Edward Elgar. ISBN 978-1-84376-331-4.

Registered once for the timing premise the method assumes rather than argues: that the deployment phase of a technological revolution is when re-derivation of operating models becomes both possible and rewarded. The market-timing argument is out of scope for this paper, which addresses the method rather than the moment.

SOCIOTECHNICAL SYSTEMS THEORY

Trist, E. L., & Bamforth, K. W. (1951). "Some Social and Psychological Consequences of the Longwall Method of Coal-Getting." *Human Relations*, 4(1).

The seminal sociotechnical study. Trist and Bamforth found that the mechanized longwall method was technically optimized and not jointly optimized, and that the loss of the social system of self-regulating work groups produced the failures the technical gains were meant to prevent. The finding is the joint-optimization principle the method inherits: the technical system and the social system must be designed together. The judgment boundary and the human-loop taxonomy are the instruments that carry it.

DECISION-RIGHTS ECONOMICS

Jensen, M. C., & Meckling, W. H. (1992). "Specific and General Knowledge, and Organizational Structure." In L. Werin & H. Wijkander (Eds.), *Contract Economics.* Basil Blackwell. Reprinted in *Journal of Applied Corporate Finance* (Fall 1995) and in Jensen, *Foundations of Organizational Strategy* (Harvard University Press, 1998).

Note this is the 1992 knowledge-and-structure paper, distinct from the authors' better-known 1976 work on agency theory. Supplies the principle that decision authority should be collocated with the knowledge required to exercise it, and the distinction that drives the principle: specific knowledge is costly to transfer among agents, general knowledge is cheap to transmit, and since no single mind can hold the knowledge relevant to all decisions, structure emerges from how decision rights are delegated. Collocation has exactly two routes, moving the knowledge to the decision rights or moving the decision rights to the knowledge. OOA operationalizes this: agents receive authority over instrumentable decisions; humans retain authority where knowledge is relational, contextual, or constitutive of the policy itself (the Trade-off Loop, 4.4). Agentic instrumentation moves the specific-general boundary that Jensen and Meckling treated as fixed by human cognition and communication cost, which is why the operating structure must be re-derived from the decisions owed rather than inherited from the constraints that no longer bind.

Sah, R. K., & Stiglitz, J. E. (1986). "The Architecture of Economic Systems: Hierarchies and Polyarchies." *American Economic Review*, 76(4), 716-727.

Supplies the error-asymmetry result the centralization choice turns on: screening structures differ in which errors they commit, and a hierarchy that requires sequential approval rejects more bad projects and more good ones alike, while a polyarchy that accepts on any single approval admits more of both. The choice between them is therefore a choice about which error is cheaper, not a choice about efficiency. OOA prices that choice rather than assuming it: the judgment boundary sets how many screens a commitment passes, and R (4.6, Appendix D.3) carries the cost of the errors each configuration lets through.

NATURALISTIC DECISION MAKING

Klein, G. (1998). *Sources of Power: How People Make Decisions.* MIT Press.
Klein, G. (2008). "Naturalistic Decision Making." *Human Factors*, 50(3), 456-460.
Klein, G. (2009). *Streetlights and Shadows.* MIT Press.
Klein, G., Snowden, D., & Chew, L. P. (2007). "Anticipatory Thinking." *Proceedings of the Eighth International NDM Conference*, Pacific Grove, CA.

Supplies the cognitive warrant for the human-owned disposition. Field expertise runs on recognition-primed pattern repertoires built through experience, and it resists replacement by rules, procedures, or analytical methods in the field settings naturalistic decision research studies. The recognition-primed model is a blend of pattern recognition and deliberate analysis:

the practitioner matches a situation to a pattern, then mentally simulates the first workable option before committing, which is why the hand-up form specified in 4.3 presents quantities with options.

Klein's streetlight-shadow distinction marks the boundary from the human side, complementing Jensen and Meckling's economic side: well-ordered decisions with stable cue structure and verifiable feedback reward proceduralization and are agent-ownable; ambiguous, time-pressured decisions defeat rule-based approaches and remain human-owned. The anticipatory-thinking work adds the seam requirement: anticipation is functional preparation to respond rather than prediction of world states, it is aimed at low-probability, high-threat events, and the documented failure mode is organizational, since individuals reliably notice weak signals while their groups reliably dismiss them. It also names the mechanism behind judgment laundering, since automated support produces passivity in the humans who oversee it. OOA draws the judgment boundary along Klein's line and engineers the seams (Phase 4) and the typed signal channel (4.3) to carry weak signals, dissent, and the conditional implications of combined events in addition to thresholded exceptions, so the seam does what the teams in the field studies failed to do.

CONDITIONS FOR INTUITIVE EXPERTISE

Kahneman, D., & Klein, G. (2009). "Conditions for Intuitive Expertise: A Failure to Disagree." *American Psychologist*, 64(6), 515-526.

Supplies the jointly authored boundary of trustworthy intuitive expertise, produced as an adversarial collaboration between the heuristics-and-biases tradition and naturalistic decision making, on the shared Simon foundation that intuition is recognition. The result is a two-condition specification, with both conditions necessary: skilled intuition develops only in environments of sufficiently high validity, where stable and learnable regularities connect observable cues to outcomes, and only where the practitioner has adequate opportunity to learn those regularities through prolonged practice with prompt, unambiguous feedback.

Two corollaries carry the operational weight. Subjective confidence is produced by fluency and coherence rather than by accuracy, so sincere professionals develop strong intuitions in low-validity environments and their confidence carries no evidential weight. And expertise fractionates by task rather than by person, so the same professional holds genuine skill on some judgments in a domain and pseudo-skill on adjacent ones. OOA reads the two conditions as the cross-camp specification of agent-ownable territory: high validity plus learnable regularity is the environment in which a policy π can be stated as executable rules, which is the knowability axis (4.4) and the instrumentability precondition (8.1) restated by both schools of decision psychology at once. The confidence corollary is the cognitive-science warrant for the anti-laundering discipline (7.4, 10.3), under which engaged exercise of human judgment is measured through decision distributions rather than trusted through attestation. Fractionation warrants the unit of analysis: skill attaches to tasks, so disposition is assigned at the decision loop (3.2), never at the

role. The paper's edge cuts in both directions, and the framework stands on the correct side of it deliberately: in low-validity environments no party, human or agent, holds predictive authority, so OOA's human-owned dispositions are warranted by non-optimizable decision variables, relational knowledge, accountability, and authority over the policy itself, and never by a claim that the expert knows.

SENSE-MAKING AND CONTEXT

Snowden, D. J., & Boone, M. E. (2007). "A Leader's Framework for Decision Making." *Harvard Business Review*, 85(11), 68-76.
Kurtz, C. F., & Snowden, D. J. (2003). "The New Dynamics of Strategy." *IBM Systems Journal*, 42(3), 462-483.

Supplies the sorting layer for disposition assignment. Cynefin partitions decisions into five contexts by how knowable cause and effect is, clear, complicated, complex, chaotic, and disorder, each carrying its own action grammar: sense-categorize-respond, sense-analyze-respond, probe-sense-respond, act-sense-respond. Clear-domain decisions are agent-owned by construction. Complicated-domain decisions are the contested zone where instrumentation advances, and the expectation that most of the portfolio's yield lives there is an expectation the ρ ledger will test across engagements through H4's output distribution rather than a finding the framework asserts. Complex-domain decisions are constitutively human-owned, because probe-sense-respond is an exercise of human judgment on patterns visible only in retrospect. Snowden's canonical executive failure, treating the complex as complicated and imposing best practice where only emergent practice exists, is OOA's mis-derivation failure stated in complexity vocabulary. Disorder names the default-to-habit risk, which for an agentically instrumented enterprise becomes default-to-automation. Cynefin's constraint taxonomy also maps onto the Theory of Constraints, its fixed and governing constraints corresponding to constraints in the constraint-theoretic sense, which is part of why the two grounding theories cohere.

CONTINUOUS DECISION LOOPS

Boyd, J. R. (2018). *A Discourse on Winning and Losing* (G. T. Hammond, Ed.). Air University Press.
Osinga, F. P. B. (2007). *Science, Strategy and War.* Routledge.

Boyd never published the loop formally; it survives in his briefings and in secondary treatments, which is the honest provenance to cite. Original briefings chiefly circulated in 1987 and include *Organic Design for Command and Control* and *The Essence of Winning and Losing.* Supplies the language of continuous decision loops (observe-orient-decide-act) and the competitive logic of loop tempo. Boyd's mature sketch is nonlinear: orientation shapes observation, decision, and action through implicit guidance and control, so most competent action bypasses explicit decision entirely, which is Klein's recognition-primed model rendered in strategic terms. Agents run explicit loops fast and shallow; the human contribution at the boundary is orientation itself, the destruction and creation of the concepts the loop runs on. OOA's fourth foundation shift

generalizes the OODA premise from adversarial maneuver to enterprise operations: the adversary is environmental drift, and tempo advantage becomes the rate at which the operating structure re-derives itself before its orientation goes stale.

MODEL-RISK GOVERNANCE

Board of Governors of the Federal Reserve System & Office of the Comptroller of the Currency (2011). *Supervisory Guidance on Model Risk Management* (SR 11-7 / OCC 2011-12).

Supplies the governance grammar for systems that decide: validation before deployment, ongoing monitoring, drift detection, documented limitations, and effective challenge. OOA's run-time control catalog (Section 10) adapts this grammar from statistical models to autonomous agents that commit rather than merely score, carrying it into boundary enforcement, drift detection, revalidation on a fixed clock, and the kill switch with its rehearsed reversion path.

STRUCTURE-MIRRORING

Conway, M. E. (1968, April). "How Do Committees Invent?" *Datamation*, 14(4), 28-31.

The source of Conway's Law: systems tend to mirror the communication structures of the organizations that build them. OOA generalizes it from software architecture to the architecture of work, which is to say that organizations ship their constraints, and uses it to predict the central failure of the retrofit response, exposed to test through the forcing-reason audit (10.1) and H2.

NEAR-DECOMPOSABILITY AND MODULAR STRUCTURE

Simon, H. A. (1962). "The Architecture of Complexity." *Proceedings of the American Philosophical Society*, 106(6), 467-482. Later collected in Simon, *The Sciences of the Artificial* (MIT Press, 1969).

Supplies the structural theory beneath seam engineering: complex systems that persist and evolve are near-decomposable, coupled strongly within subsystems and weakly across them, so that the short-run behavior of each subsystem is approximately independent of the others and its long-run behavior depends on them only in aggregate. The watchmakers Hora and Tempus supply the evolutionary corollary, that a system built from stable intermediate subassemblies survives interruption and reassembles at a fraction of the cost of one that must be completed whole, which is why surviving complexity is modular and why reversion paths (6.5) are specifiable at all.

OOA installs seams (4.5) at exactly these near-decomposable interfaces: a seam is admissible where the coupling across it is weak enough that the flow on either side can be re-derived, encapsulated, and instrumented without the counterparty's internal structure entering the design, and the encapsulate-behind-an-interface seam pattern is near-decomposability used as a construction rule rather than observed as a fact. The same property licenses the decision loop as the unit of analysis (3.2), held fixed and analyzed short-run because it is weakly coupled to its surroundings, its five-tuple the aggregate description that survives while the rest is re-derived.

Simon's distinction between formal hierarchy, the authority relation drawn on the org chart, and hierarchy in his broad sense, any nested near-decomposable structure, is the distinction the method depends on: OOA rejects the former as foundation-contaminated (3.2) while building its seam architecture on the latter, so the citation grounds the modularity without endorsing the org-chart nesting the re-derivation exists to dissolve.

REGISTER NOTE · CONVERGENCE

The five sources triangulate one claim from three disciplines and two rival schools of a fourth. Organizational economics (Jensen and Meckling, with Sah and Stiglitz on the error asymmetry) shows decision authority follows knowledge. Cognitive field research (Klein) shows which knowledge resists transfer and why. Complexity science (Snowden, with Boyd as strategic ancestor) shows how to sort decisions by the knowability of their cause-and-effect structure. Decision psychology, speaking through both of its traditions at once (Kahneman and Klein), states the environmental conditions under which proceduralizable skill can exist at all. None of the five treats the resulting boundary as a designed, economically evaluated artifact. OOA supplies that step through the disposition taxonomy and the yield formula $\rho = V / (C + R)$.